



ICR A PARTIR DE TRAÇOS D'ESCRITURA NATURAL CAPTADES AMB UN BOLÍGRAF DIGITAL

Memòria del projecte
d'Enginyeria en Informàtica
realitzat per

Ivan Navarro Parreño

i dirigit per

Gemma Sanchez Albaladejo

Bellaterra, 1 de Setembre de 2009

El sotasignat, Gemma Sanchez Albaladejo
Professor/a de l'Escola Tècnica Superior d'Enginyeria de la UAB,

CERTIFICA:

Que el treball a què correspon aquesta memòria ha estat realitzat sota la seva direcció per en Ivan Navarro Parreño.

I per tal que consti firma la present.

Signat: Gemma Sanchez Albaladejo
Bellaterra, 1 de Setembre de 2009

El sotasignat, Bernat Ballester
de l'empresa, PlastiaSite S.A.,

CERTIFICA:

Que el treball a què correspon aquesta memòria ha estat realitzat en l'empresa sota la seva supervisió mitjançant conveni firmat amb la Universitat Autònoma de Barcelona.

Així mateix, l'empresa en té coneixement i dóna el vist-i-plau al contingut que es detalla en aquesta memòria.

Signat: Bernat Ballester

Bellaterra, 1 de Setembre de 2009

*Als meus pares,
perque m'ho heu donat tot.
A la Rosa,
per estar al meu costat en tot moment.*

Agraïments

En primer lloc, gràcies per la oportunitat que m’ha dona PlastiaSite S.A. de realitzar aquest treball, especialment al Bernat Ballester i al seu grup (Albert, Iván, Ismael, Raquel, Doris, Carles, Xavier i Miguel). I a la Gemma Sánchez per la seva bona guia i direcció durant tot el desenvolupament del projecte.

Gràcies als companys i amics de la Universitat (Víctor, Rocio, Irene, Peña, Miguel, Uri, Aroa, Guillem, etc...). Gràcies pels molts i grans moments que hem tingut en aquesta etapa de la nostra vida. Tots aquests anys hem tingut inoblidables hores de rialles, des de les estones de descans i bar, fins a les matinades inhumanes a la biblioteca.

Als meus amics, en especial a Paco, Sami i Xavi, que han tingut la paciència de no deixar-me de trucar, tot i no fer-los el cas que es mereixien, durant aquests últims anys.

Finalment, i precisament el més important, gràcies a la meva família. Tota la família (que no som pocs). Per donar-me l’oportunitat i recolzar-me en aconseguir aquesta fita i d’altres al llarg de la meva vida. Molt especialment a la Rosa, - *“La teva paciència és infinita” ;-)* .

Índex

1	Introducció	1
1.1	Motivació i objectius	2
1.2	Estudi de viabilitat del projecte	3
1.2.1	Utilització d'un bolígraf digital	3
1.2.2	Reconeixedors de caràcters	3
1.3	Planificació del projecte	4
1.3.1	Planificació inicial del projecte	4
1.3.2	Planificació final del projecte	6
1.3.3	Pressupost	7
1.4	Contingut de la memòria	8
2	Estat de l'art	9
2.1	Aparició dels primers reconeixedors de caràcters	10
2.2	Diferents tecnologies: OMR, OCR i ICR	10
2.3	Reconeixement Off-line i On-line	11
2.4	El bolígraf digital i altres dispositius	13
2.5	Tecnologia a fer servir	13
3	Descripció del dispositiu i els ICR escollits	15
3.1	Bolígraf digital de Logitech® i eines d'Anoto™	16
3.1.1	Disseny i impressió del paper intel·ligent	17
3.1.2	Bolígraf digital Logitech® IO™ 2	19
3.1.3	Comunicació bolígraf-aplicacions	20

3.1.4	Aplicacions	20
3.2	Descripció i elecció dels reconeixadors ICR	21
3.2.1	Reconeixedor Heloise Hse	21
3.2.2	Reconeixedor de Microsoft®	23
4	Disseny de l'entorn d'anàlisi	25
4.1	Objectiu de l'anàlisi	26
4.2	Entorn de proves	26
4.2.1	Presa de contacte: primeres proves amb el bolígraf digital i l'adquisició de dades	26
4.2.2	Primera fase: comunicació amb els reconeixadors i primers resultats	27
4.2.3	Segona fase: implementació amb base de dades externa i avaluació Precision-Recall	29
4.2.4	Tercera fase: combinació de reconeixadors	32
4.3	Bases de dades	33
4.3.1	Base de dades pròpia	33
4.3.2	Base de dades externa	34
5	Avaluació dels sistemes de reconeixement	37
5.1	Tècniques d'anàlisi de resultats	38
5.2	Avaluació dels resultats	39
5.2.1	Reconeixadors de Microsoft® i Heloise Hse amb base de dades pròpia	39
5.2.2	Reconeixadors de Microsoft® i Heloise amb base de dades externa	41
5.2.3	Combinació dels reconeixadors	44
6	Conclusions	49
6.1	Objectius assolits	50
6.2	Línies futures	51

7	Annexes	53
7.1	Requeriments del sistema	53
7.2	Flux d'execució i manipulació amb solucions Anoto TM	54
7.3	Base de dades pròpia	55
7.4	Disseny de les diferents fases	57
7.4.1	Disseny a la primera fase	57
7.4.2	Disseny a la segona fase	58
7.4.3	Disseny a la tercera fase	59
7.5	Taules Precision-Recall	60
7.6	Matrius de Confusió dels resultats	61

Índex de figures

1.1	Planificació inicial	5
1.2	Planificació final	6
3.1	Esquema de passos per la creació de formularis Anoto™	17
3.2	Exemple del patró d'Anoto™	18
3.3	Segmentació del patró d'Anoto™	18
3.4	Esquemes del bolígraf digital	19
3.5	Proves del reconeixedor Heloise Hse amb figures.	22
4.1	Mapa relacional de la base de dades	33
5.1	Esquema Precision-Recall	39
5.2	Comparativa de nombres encertats en fase 1	41
5.3	Comparativa de nombres encertats en fase 2	42
5.4	Comparació entre reconeixadors obtinguent una sola alternativa a la segona fase	43
5.5	Comparació entre reconeixadors dependents d'escriptor	46
5.6	Comparació entre reconeixadors independents d'escriptor	46
5.7	Comparació entre reconeixadors a la fase 3	47
7.1	Flux d'execució i manipulació amb solucions Anoto™	54
7.2	Esquema de la creació dels fitxers XML	57
7.3	Esquema del flux d'execució i de les dades a l'entorn de test.	58
7.4	Esquema de la millora del reconeixedor mixte	59
7.5	Matriu de confusió reconeixedor de Microsoft®	62

7.6	Matriu de confusió reconeixedor de l'Heloise Hse (Independent d'escriptor)	63
7.7	Matriu de confusió reconeixedor de l'Heloise Hse (Dependent d'escriptor)	64
7.8	Matriu de confusió reconeixedor mixte (Independent d'escriptor)	65
7.9	Matriu de confusió reconeixedor mixte (Dependent d'escriptor) .	66

Índex de taules

5.1	Comparativa entre els resultats obtinguts a la primera fase	40
5.2	Comparativa entre els resultats obtinguts a la segona fase. Re- coneixedors dependents d'escriptor	42
5.3	Comparativa entre els resultats obtinguts a la segona fase. Re- coneixedors independents d'escriptor	42
5.4	Comparativa de resultats dels reconeixedors configurats per re- tornar una sola alternativa	43
5.5	Valors de l'exemple del funcionament del reconeixedor mixte . . .	44
5.6	Percentatges d'encert amb el reconeixedor mixte	45
7.1	Taula Precision-Recall associada al reconeixedor de Microsoft® .	60
7.2	Taula Precision-Recall associada al reconeixedor Heloise Hse (In- dependent d'escriptor + B.D. Externa)	60
7.3	Taula Precision-Recall associada al reconeixedor Heloise Hse (De- pendent d'escriptor + B.D. Externa)	61

Capítol 1

Introducció

Aquest primer capítol és la introducció al desenvolupament i anàlisi del treball realitzat. A continuació s'exposen els punts que contindrà, i s'introduirà al lector en aquest projecte. Les diferents seccions son:

- Motivació i objectius:

Una breu explicació dels objectius marcats i la finalitat del projecte.

- Estudi de viabilitat del projecte:

Breu introducció dels punts que condicionen la viabilitat del projecte.

- Planificació del Projecte:

Planificació inicial i final amb una petita explicació de les diferents tasques que l'han fet possible.

- Contingut de la memòria:

S'exposarà un breu resum de cadascun dels diferents capítols que formen la present memòria.

1.1 Motivació i objectius

A la indústria, sobretot en cadenes de producció, els operaris treballen independentment de la part administrativa. Normalment, el seguiment de les peces que es fabriquen només es pot fer quan els informes dels diferents operaris arriben físicament en paper a l'administració de l'empresa.

La possibilitat d'escorçar el temps de coneixement de l'estat d'una cadena de producció, pot fer més flexible les preses de decisió i evitar així unes solucions desfasades al temps real de dita producció. Així com augmentar l'agilitat en l'administració digitalitzant els documents.

L'objectiu del projecte és realitzar un estudi, per la posterior implementació, d'un Reconeixedor de Caràcters Intel·ligent (Intelligent Character Recognition - ICR) a partir d'un sistema que recull els traços fets en l'escriptura manuscrita de caràcters.

Aquest nivell de reconeixement formarà part d'un sistema més ampli incorporat en una cadena de producció. Aquest, té com a finalitat retransmetre, al final de la jornada laboral, l'estat de la cadena al departament administratiu.

La finalitat, més concretament, és realitzar un anàlisis entre diferents reconeixadors de traços i arribar a la conclusió de com reconèixer nombres amb la millor precisió possible. A partir d'aquests nombres s'identificaran els diferents formularis i podran ser emmagatzemats en bases de dades i manipulats informàticament.

Al finalitzar l'avaluació es decidirà si n'hi ha prou amb aquests reconeixadors, comercials avui en dia, o bé és necessària una nova implementació per obtenir uns resultats més satisfactoris.

1.2 Estudi de viabilitat del projecte

Per demostrar la viabilitat d'aquest projecte, a continuació es dona a conèixer els dispositius i tecnologies existents que ens permetran resoldre el problema plantejat.

1.2.1 Utilització d'un bolígraf digital

Com a mètode d'entrada de dades s'utilitzarà el bolígraf digital (Logitech® IO™2 DigitalPen)[5]. Aquest és un dispositiu que ens permet emmagatzemar les dades que recull mentre un usuari escriu sobre un formulari.

Els formularis utilitzats per aquest bolígraf son completament manipulables gràcies a Anoto™[6], una solució *software* que enfoca els bolígrafs digitals a una personalització segons les necessitats.

El dispositiu és prou robust com per ser utilitzat en entorns hostils i bruts com pot ser un taller mecanitzat. Tot i que existeixen altres dispositius que podrien tenir la mateixa funcionalitat, com les PDAs, el bolígraf digital comporta un impacte mínim en els treballadors. Actualment, ja omplen formularis de forma manuscrita al pas dels diferents estats de la producció.

1.2.2 Reconeixadors de caràcters

Existeixen molts i diferents sistemes capaços de reconèixer caràcters i símbols, així com diferents punts de vista que intenten resoldre la mateixa problemàtica. Per aquest motiu, agafarem com a punt de referència dos implementacions ja conegudes en aquest àmbit: el reconeixedor de caràcters de Microsoft®[7, 8] i el de Heloise Hse[9] provinent de la Universitat de Berkeley a Califòrnia.

1.3 Planificació del projecte

La planificació d'aquest projecte ha estat en base a un temps acotat per unes dates d'inici i finalització. Es diferencien una planificació inicial i una de final ja que s'han realitzat rectificacions del projecte segons les condicions que han sorgit al llarg de l'anàlisi.

Les planificacions estan dividides en tasques representades amb gràfiques de GANTT per una millor visualització temporal de la durada de cadascuna. Els temps de dedicació estan disposats en setmanes i amb una dedicació diària de 4 hores. Només es tenen en compte les jornades laborals i no els períodes vacacionals o caps de setmana.

Finalment, es mostra el pressupost que comporta realitzar un estudi com el present.

1.3.1 Planificació inicial del projecte

Al començament del projecte hi havien punts o tasques que eren clars i necessaris per una finalització satisfactòria. Així podem destacar un estudi dels requeriments del projecte, estudi de l'estat de l'art, un desenvolupament de l'entorn d'avaluació i la pròpia avaluació dels resultats obtinguts.

D'una altra banda, les especificacions de les tasques de desenvolupament i test no eren tan clares degut a la dependència dels resultats obtinguts. Per tant, el diagrama de la figura 1.1 mostra una planificació poc definida però que marca els temps d'inici i finalització aproximats de cada apartat respectant les dates d'entrega.

Les diferents tasques previstes han estat les següents:

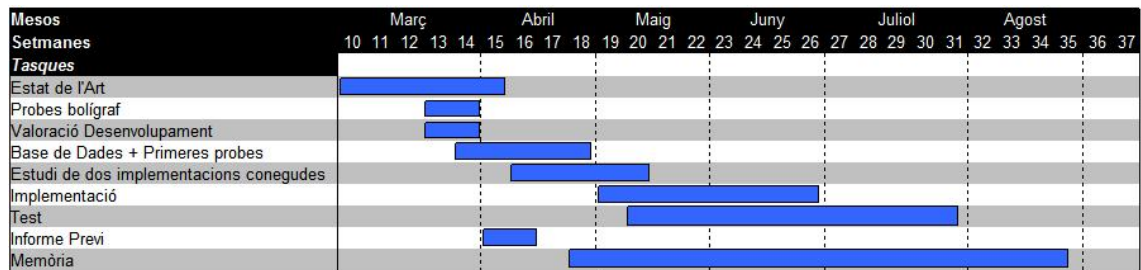


Figura 1.1: Diagrama de Gantt que mostra la planificació inicial del projecte.

- **Estat de l'art:** reflexa el temps necessari per recollir prou informació que ens introdueixi en el món del reconeixement de caràcters, concretament de nombres, manuscrits on-line.
- **Proves bolígraf i Valoració desenvolupament:** representen dos punts amb els quals es pren el primer contacte amb el bolígraf digital, l'entorn de desenvolupament dels formularis i les respectives APIs.
- **Base de Dades i primeres proves:** marquen el primer desenvolupament i joc amb les eines que ofereixen AnotoTM i Logitech®.
- **Estudi de dos implementacions conegudes:** període de cerca i estudi d'almenys dos reconeixadors coneguts en aquest àmbit pels seus bons resultats.
- **Implementació:** desenvolupament de l'entorn d'avaluació dels reconeixadors. Aquest és imprecís degut a la seva evolució segons els resultats obtinguts.
- **Test:** marca el temps d'anàlisi dels resultats obtinguts.
- **Informe previ i Memòria:** temps en que es portarà a terme l'escriptura dels informes.

1.3.2 Planificació final del projecte

Finalment, el projecte ha estat realitzat dins els terminis marcats. Be, no solament el projecte en conjunt, sinó també cadascuna de les tasques proposades. En concret destacar, ara sí, la especificació de les implementacions i els tests de l'entorn d'avaluació. Veure figura 1.2.

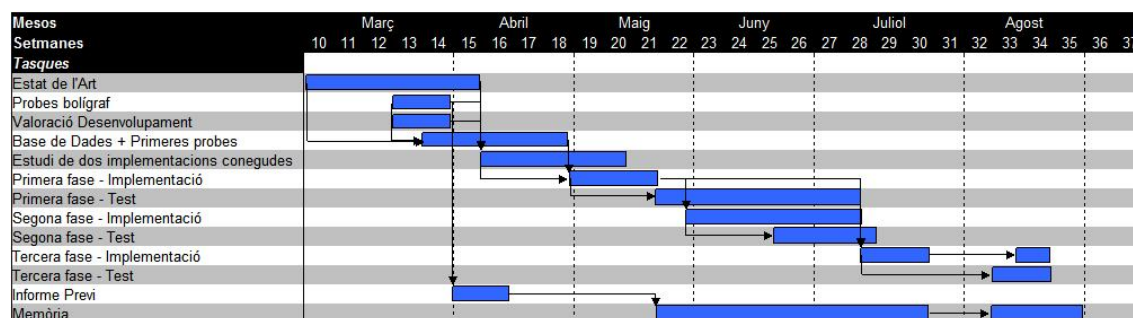


Figura 1.2: *Diagrama de Gantt que mostra el temps amb el que s'ha dut a terme del projecte.*

El desenvolupament es divideix en tres fases que inclouen la implementació i test de l'entorn d'avaluació:

- **Primera fase:** recull de les primeres mostres amb el bolígraf digital i implementació de l'entorn d'anàlisi amb una base de dades pròpia amb la que s'avalua cadascun dels reconeixadors per separat. Tests amb matrius de confusió dels resultats.
- **Segona fase:** implementació i test d'un entrenament amb una base de dades externa. Tests amb matrius de confusió i anàlisi precision-recall.
- **Tercera fase:** un cop analitzats els reconeixadors per separat, implementació de reconeixadors combinats en busca de millors resultats. Test amb matrius de confusió.

En total, aquest projecte ha estat efectuat en un total de 480 hores (aproximadament).

1.3.3 Pressupost

La realització d'aquest estudi comporta unes despeses de maquinari, programari i hores justificades en investigació, desenvolupament i avaluació de resultats, que seran esquematitzades a continuació:

És necessari un sistema informàtic personal (**1.200 Euros**) per al desenvolupament de l'entorn d'anàlisi.

El bolígraf digital té un preu de 130 Euros aproximadament, tot i que per fer proves i desenvolupament, AnotoTM posa a disposició un kit de demostració amb un cost de **500 Euros**. Aquest conjunt incorpora el necessari per una utilització més personalitzada.[5, 6]

Altres eines com les de programació o reconeixement són gratuïtes. El projecte es basa en els llenguatges Java i C# amb unes APIs que s'ofereixen al kit de demostració. Respecte als reconeixadors, tal i com s'esmentarà en següents seccions, es farà servir el proporcionat per Microsoft® (gratuït per usuaris de Windows) i l'Heloise Hse (de codi obert).[7, 8, 9]

Si tenim en compte la feina portada a terme per un Enginyer Informàtic Superior amb titulació i les hores dutes a terme podem estimar un pressupost de "mà d'obra" de $480h * 2.500\text{Eur}/\text{mes} * 1\text{mes}/160\text{hores} = \mathbf{7.500\text{ Euros}}$.

Tot plegat, les despeses de desenvolupament i avaluació de l'anàlisi tindria un preu aproximat de 9.200Euros. Si incrementem aquest preu amb un benefici del 200% obtindríem un preu final de producte de **18.400 Euros**.

1.4 Contingut de la memòria

La memòria esta dividida en sis capítols. El present, la **“Introducció”**, situa al lector de forma breu en el contingut del document. Seguidament, a l’**“Estat de l’art”** es mostra l’estat en que es troba el sector del reconeixement de documents.

El capítol **“Descripció dels dispositius i els ICR escollits”** mostrarà quines de les tecnologies existents es faran servir i el motiu de les diferents eleccions.

Als capítols **“Disseny de l’entorn d’anàlisi”** i **“Avaluació dels sistemes de reconeixement”**, s’explica, de forma cronològica, el procés de desenvolupament i l’avaluació dels resultats.

Finalment, al capítol **“Conclusions”**, es farà l’avaluació de la realització d’aquest projecte i s’exposaran les conclusions arribades a partir dels resultats obtinguts.

Es tancarà el document amb els **“Annexes”** i **“Bibliografia”** pertinents.

Capítol 2

Estat de l'art

A continuació es farà un recorregut pels inicis de la història en tecnologia de reconeixement de caràcters. Veurem que es tracta d'un àmbit molt estudiat i amb un ventall ampli de solucions. Les diferents seccions mostraran:

- Un breu apunt dels primers indicis dels reconeixadors.
- Els diferents tipus de reconeixadors segons la seva naturalesa (marques, caràcters i altres).
- Diferenciació entre el reconeixement de símbols a partir d'imatges, o tenint en compte altres variables, i els algorismes relacionats.
- Existència dels diferents dispositius al mercat que poden servir com a punt d'adquisició de dades.
- Finalment, les preses de decisió per escollir els diferents components i tecnologies.

2.1 Aparició dels primers reconeixadors de caràcters

La idea que una màquina pugui tenir la capacitat de reconèixer caràcters no és nova. Els primers vestigis daten del 1929 quan Gustav Tauschek obté la primera patent a Alemanya i als Estats Units mitjançant fotodetectors[1].

En aquest temps han sorgit diferents projectes mirant de resoldre el problema de reconèixer caràcters o textos, ja siguin impresos en fulls com recollits a partir de l'escriptura manual. Podem nombrar per exemple, Xstroke, wayV, Cell-Writer, Xmerlin o Rosetta[10] distingits al món del codi lliure, així com Tesseract i OCRopus de Google[11]; o bé Microsoft® Office Document Imaging i Office OneNote[12] de caire comercial.

Actualment, el reconeixement de caràcters és considerat un problema sobradament resolt, tot i que en certes situacions es requereixen uns percentatges d'error menors dels aconseguits fins ara. Té una especial complexitat el reconeixement de caràcters en documents on no està clar la identificació del text per proximitat a àrees d'interferències com dibuixos o altres tipus de soroll. Així com la digitalització de la gran quantitat de documents històrics per la seva antiguitat i varietat lingüística.

2.2 Diferents tecnologies: OMR, OCR i ICR

Avui en dia podem distingir diferents conceptes i/o mètodes pel reconeixement de caràcters i símbols. Ens trobem amb tres tecnologies diferents[2]:

- **Reconeixement de marques òptic (OMR):**

Detecta la presència o manca d'una marca, però no la seva forma. Un

exemple d'aquesta tecnologia son els formularis que es troben als tests o enquestes.

- **Reconeixement de caràcters òptic (OCR):**

Proporciona la capacitat d'interpretar els diferents símbols a partir de l'escaneig d'imatges estàtiques com poden ser els caràcters que ens trobem en impresos o textos d'impremta.

- **Reconeixement de caràcters intel·ligent (ICR):**

Proporciona la capacitat d'interpretar els caràcters manuscrits i/o textos en cursiva mitjançant una computadora.

Parlem de caràcters manuscrits a aquells que han estat escrits manualment amb un bolígraf i aïllats els uns dels altres. Per una altra banda, els textos en cursiva son aquells que son escrits a mà lligant unes lletres amb altres sense deixar espais entre elles formant paraules.

2.3 Reconeixement Off-line i On-line

Els caràcters manuscrits poden ser adquirits, emmagatzemats i interpretats de dos formes diferents: Off-line i On-line.

El reconeixement Off-line[3] es centra en l'anàlisi de símbols a partir d'imatges, per exemple després d'escanejar un paper i obtenint els diferents caràcters en escala de grisos. Mitjançant múltiples algorismes sobre aquestes imatges es poden separar per classes, per exemple, i adquirir el valor de forma analítica que nosaltres reconeixem de forma natural.

Per una altra banda, el reconeixement On-line[3, 4] no depèn d'una simple imatge, sinó d'un conjunt de dades, tals com punts, marques de temps, traços,

nivell de pressió; que s'han pogut recollir al mateix moment de realitzar l'escriptura, normalment manuscrita o en cursiva.

Gràcies a diferents tècniques es poden analitzar aquestes dades i determinar la relació amb els caràcters d'una forma més acurada i gairebé a temps real.

Alguns dels algorismes relacionats al reconeixement de patrons en general, i de caràcters en particular, serien els següents:

- **Hidden Markov Model (HMM)[13]:** És un model estadístic que tracta el reconeixement dels patrons a partir d'uns valors probabilístics. En cas de tenir més d'una opció a escollir, el sistema selecciona el següent estat a partir d'aquests valors prèviament especificats.
- **K-Nearest Neighbor (K-NN)[14, 18]:** Consisteix en convertir els punts que representen els traços en vectors amb direccions aproximades. La mostra pertanyerà al conjunt de mostres més nombroses dins dels subconjunt dels K elements més “propers”.
- **Neural Networks (NN)[15]:** Aprofita les capacitats de classificació d'aquesta tecnologia per classificar.
- **Minimum Distance (MD)[18, 19]:** És una tècnica empleada d'una forma similar a K-Nearest Neighbor. El símbol a reconèixer serà membre del grup més proper segons un conjunt de característiques, mitjançant el càlcul de la distància euclidiana normalitzada.
- **Support Vector Machine (SVM)[18, 16, 17]:** Classificador multiclases. Separa les dades a classificar en diferents conjunts sobre un espai n-dimensional segons un hiperplà. Escollint les mostres que entre conjunts hi hagi el màxim espai possible.

2.4 El bolígraf digital i altres dispositius

Existeixen diferents dispositius útils per introduir informació, com caràcters o nombres, a un ordinador. Des de les taules digitalitzadores o escàners òptics fins a computadores personals amb pantalles tàctils o bolígrafs digitals; els diferents sistemes que ens permeten utilitzar aquestes tecnologies és àmpliament nombrosa, cada dia més.

Per decidir quin d'aquests dispositius és més adient per a la resolució del problema s'ha de pensar en les característiques i necessitats de l'usuari final. En concret, aquest punt es resol per requeriment del sistema.

2.5 Tecnologia a fer servir

S'utilitzarà un bolígraf digital de la marca Logitech®. Aquest bolígraf té la capacitat de recollir els traços, formats per punts i marques de temps, que l'usuari fa al escriure sobre uns formularis concrets.

Aquestes dades es poden analitzar i, aplicant-hi les tecnologies esmentades anteriorment, podem automatitzar un reconeixement sobre les dades que l'usuari escriu en un temps gairebé real.

Altrament dit, el reconeixement d'aquests nombres el farem amb una tecnologia ICR i, per tant, un reconeixement On-line sobre les dades adquirides amb el bolígraf digital.

Capítol 3

Descripció del dispositiu i els ICR escollits

En aquest capítol s'examinaran a fons les característiques del bolígraf digital, així com els entorns de desenvolupament que ofereixen Logitech® i Anoto™. Tant mateix, s'exposarà una descripció dels reconeixadors de Microsoft® i Heloise Hse, i els punts de decisió que han definit l'elecció d'aquests.

3.1 Bolígraf digital de Logitech® i eines d'Anoto™

El tàndem format pel bolígraf digital Logitech® IO™2 i les eines de formularis d'Anoto™ és interessant des del punt de vista productiu. El servei que ofereixen consta de la possibilitat de creació de formularis personalitzats i una fàcil adquisició i posterior tractament de les dades manuscrites.

El sistema és fàcilment separable en quatre punts dins de tot el procés, des de la creació d'un formulari fins al tractament de les dades. Veure figures 3.1 i 7.1:

- Disseny i impressió del paper intel·ligent:

Creació d'un paper "inel·ligent" on es disposa el formulari desitjat i fa possible la orientació del bolígraf digital en aquest.[6, 20]

- Bolígraf digital Logitech® IO™ 2:

El dispositiu que fa possible l'adquisició de dades manuscrites.[5, 20]

- Comunicació bolígraf-aplicacions:

És la fase encarregada d'adquirir les dades recollides pel dispositiu i transportar-les segons les necessitats[6, 20]. A la figura 3.1, aquest punt correspon a les seccions de connexió, *routing* i transport.

- Aplicacions:

Fase final de tot el procés on es tracten les dades recollides segons les finalitats.[6, 20]

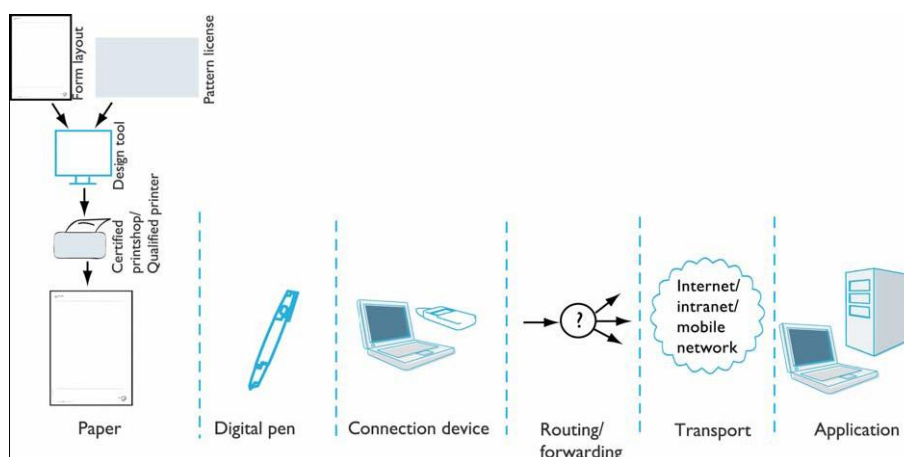


Figura 3.1: Aquest esquema mostra les diferents parts en que es divideix el procés de creació de formularis, recollida de traços, transport de les dades i us, i creació d'aplicacions.

3.1.1 Disseny i impressió del paper intel·ligent

La creació de formularis personalitzats és una de les claus del sistema. Anoto™ dona la facilitat de realitzar aquest procés sobre uns fulls que s'anomenen “intel·ligents” ja que permetran un reconeixement i informatització de les dades escrites a mà.

Els formularis tenen la peculiaritat d'estar impresos juntament amb un patró de punts que ocupa tota la superfície del full i situa a un bolígraf digital on s'està escrivint. Es tracta d'una malla quadrada de punts minúsculs (100 micròmetres de diàmetre) separats entre ells per una distància de 0.3 mm. Veure la figura 3.2.

En regions quadrades de 6x6 punts (cadascun en diferent posició) es pot donar lloc a enormes àrees sense repetició de combinacions d'aquests punts. Aquesta combinació podria arribar a cobrir un àrea de 60 milions de kilòmetres quadrats. Aquesta excedeix el conjunt dels continents Europeu i Asiàtic junts.

Per administrar aquesta gran quantitat de superfície, Anoto™ la divideix en

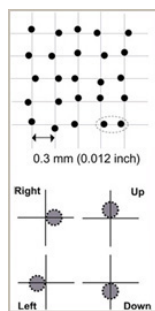


Figura 3.2: Aquesta és una mostra del patró d'AnotoTM.

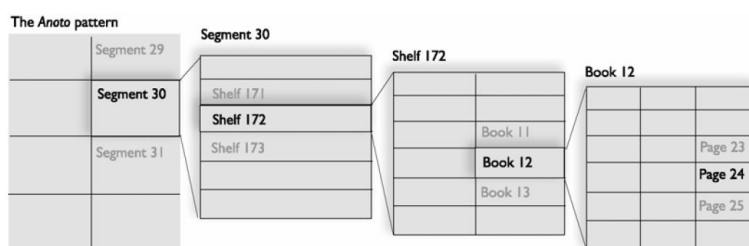


Figura 3.3: Aquesta figura mostra com es divideix el patró d'AnotoTM. Degut a la seva grandària, s'ha de dividir en diferents parts i enumerar-les com si direccions IP es tractessin.

diferents parts separades per punts com si una direcció IP es tractés: “*Segments*”, “*Shelves*”, “*Books*” i “*Pages*”. Una possible identificació de la part del patró sobre la que s’ha imprès el formulari seria: 4.3.2.1 (segment 4, prestatge 3, llibre 2 i pàgina 1). Veure la figura 3.3.

Inserides als formularis, es poden especificar diferents àrees destinades a diferents propòsits. Així, podem trobar àrees d’usuari, específiques per escriure textos sobre elles o dibuixar; i *pidgets*, petites àrees amb diferents funcionalitats com canviar el color o tamany del traç que s’efectuarà.

Totes aquestes possibilitats i d’altres s’especifiquen a l’hora de crear el formulari, i s’emmagatzemen en un arxiu amb extensió *.PAD (Paper Application Definition).

3.1.2 Bolígraf digital Logitech® IO™ 2

El bolígraf digital és un dispositiu amb una forma molt similar (una mica més gros) a un bolígraf convencional. Existeixen diverses marques i models compatibles amb aquest sistema. En concret, es farà servir el bolígraf de Logitech® IO™ 2.

Per saber en quin punt del paper s'està efectuant l'escriptura, el bolígraf digital pren captures amb una micro-càmera (entre 50 i 100 captures per segon) de l'àrea quadrada que té just a sota de la punta.

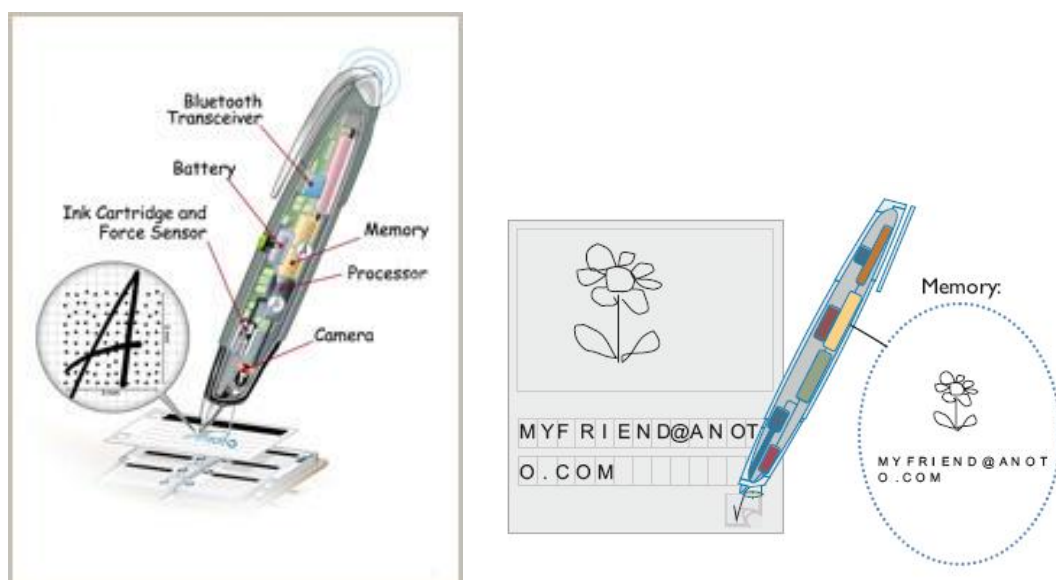


Figura 3.4: Aquest esquema mostra les parts principals del bolígraf digital i el contingut que es guarda a memòria després d'escriure sobre un formulari prèviament configurat.

Dins del bolígraf digital, una petita memòria emmagatzema els traços que l'usuari efectua sobre el formulari per posteriorment ser retransmesos. Veure figura 3.4.

3.1.3 Comunicació bolígraf-aplicacions

Un cop l'usuari ha escrit amb el bolígraf digital i es disposa a treballar amb les dades emmagatzemades a la memòria del mateix, l'únic que ha de fer és dipositar aquest sobre la seva base connectada alhora a un ordinador.

Aquesta mateixa base serveix tant per descarregar la informació del bolígraf com per carregar la seva bateria. Quan dita informació es descarregui correctament, ho farà en forma de fitxer amb extensió *.PGC (Pen Generated Coordinates).

El binomi d'arxius *.PAD i *.PGC poden ser exportats i compartits mitjançant diverses vies (internet, intranet, xarxes de bluetooth) amb la finalitat de ser tractats per les aplicacions desenvolupades per tal propòsit.

3.1.4 Aplicacions

Aquesta és la última fase i una de les més importants de tot el procés d'escriptura amb aquesta tecnologia. Les aplicacions, ja siguin d'empreses externes o be desenvolupades pròpies, seran les destinatàries dels arxius *.PAD i *.PGC.

AnotoTM proporciona una API Java amb la qual es poden extreure els diferents traços realitzats amb el bolígraf digital i tractar-los segons convingui.

Les diferents opcions poden ser adquirir només els punts i marques de temps que formen els traços, transformar aquests en imatges de diferents formats (PNG, JPG, BMP,...), o be efectuar les accions que han estat especificades mitjançant els *pidgets* al moment d'escriure.

3.2 Descripció i elecció dels reconeixadors ICR

Davant l'elecció dels reconeixadors, s'ha de tenir en compte que els símbols que volem reconèixer són únicament nombres del 0 al 9, manuscrits, compatible amb tecnologies on-line, que faciliti el desenvolupament i una alta efectivitat, ja que s'ha de reconèixer automàticament l'identificació dels formularis.

Dins de les possibilitats, s'han escollit dos reconeixadors de caràcters coneguts:

- Reconeixedor Heloise Hse [9]
- Reconeixedor de Microsoft® [7, 8]

3.2.1 Reconeixedor Heloise Hse

El reconeixedor Heloise Hse és un reconeixedor de patrons per a models d'escriptura manuscrita, on-line i de propòsit general. Creat per Heloise Hse i A.Richard Newton del Departament d'Enginyeria Elèctrica i Ciències de la computació a la Universitat de Berkeley a Califòrnia.

Que sigui un reconeixedor de propòsit general significa que no està encarat a reconèixer cap mena de model en especial, i per tant, és l'usuari que l'ha d'entrenar per reconèixer els models necessaris, sempre i quan siguin símbols manuscrits on-line.

El caracteritza ser de codi font d'ús lliure sense restriccions, sempre i quan s'especifiquin els seus orígens a la aplicació final. Està desenvolupat en Java, proporciona una API i aplicacions ja desenvolupades que mostren el funcionament.

Conté la possibilitat de fer-se servir amb qualsevol de tres nuclis de reconeixement (Support Vector Machine - *SVM*, Minimum Distance - *MD* o K-Nearest

Neighbor - K -NN).

El propi paquet proporciona i incorpora un sistema de preprocés mitjançant moments de Zernike que varien les dades d'entrenament i test mitjançant translacions i redimensionaments per aconseguir una homogeneïtat entre les mostres. Com a resultat d'aquest procés, totes les figures seran normalitzades a mida 100x100 píxels.

La documentació de l'Heloise Hse proporciona un anàlisi realitzat amb uns models de símbols que es poden trobar en varietat d'aplicacions gràfiques així com editors UML o el propi Microsoft® PowerPoint. Les gràfiques de la figura 3.5 mostren la seva efectivitat amb tots tres nuclis.

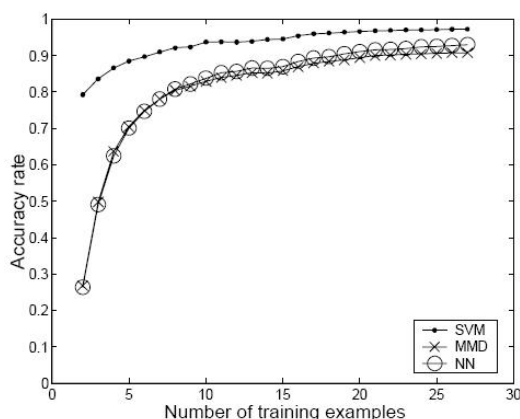
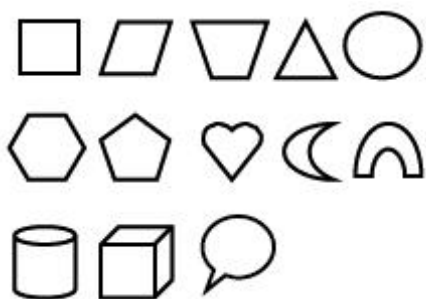


Figura 3.5: Aquesta és una mostra dels resultats obtinguts a les proves de reconeixement amb figures de diagrames. Reconeixements independents d'escriptor.

Els resultats obtinguts amb el nucli SVM son els millors de totes les avaluacions i serà el que es farà servir en l'avaluació de nombres.

3.2.2 Reconexidor de Microsoft®

El paquet del reconexidor de Microsoft® incorpora una API de diversos llenguatges i una guia amb diferents exemples, aquests donen la possibilitat a un programador d'interactuar amb aquestes eines per poder crear lliurement aplicacions que les continguin o depenguin d'elles.

Aquest entorn de desenvolupament requereix el MS .NET Framework SDK v2.0 i el Visual Studio. Com a llenguatges de programació, el desenvolupador té suport i pot escollir entre Visual C#, Visual Basic .NET o Visual Basic 6.0. L'SDK dona així l'oportunitat de descobrir les capacitats d'aquest paquet de Microsoft®, d'entre elles es poden destacar:

- Reconeixement i manipulació de tinta electrònica.
- Eines de reconeixement de patrons escrits amb tinta electrònica.
- Transport i manipulació de tinta electrònica mitjançant xarxes com Internet.

La tinta electrònica no és més que el conjunt en traços o *strokes* amb punts de coordenades (X,Y) i és tractat com objectes en els llenguatges de programació.

El reconexidor de Microsoft® és, a la vista del programador o l'usuari, una caixa negra que, especificant les opcions o paràmetres adequats i uns traços en forma d'objecte, retorna els resultats del reconeixement.

La tecnologia utilitzada per implementar-ho són les xarxes neurals amb temps de retard (Time Delay Neural Network). Un punt important és la capacitat de reconèixer textos manuscrits on-line en alguns idiomes (Xinès (Tradicional), Xinès (Simplificat), Anglès americà, Anglès (U.K.), Francès, Alemany, italià, Japonès, Coreà i Espanyol).

L'aplicació d'aquesta tecnologia en les diferents solucions del gegant informàtic i la implantació a Windows XP, Vista i el seu successor (Windows 7), dona bones referències de la capacitat de reconeixement del paquet.

Capítol 4

Disseny de l'entorn d'anàlisi

Aquest capítol reflectirà les diferents fases que s'han dut a terme en la realització d'aquest projecte.

Es divideix el contingut en dos parts importants. Per una banda, el desenvolupament de l'entorn d'anàlisi especificant les diferents fases que s'han dut a terme. Per una altra banda, la creació i us de les bases de dades que s'han fet servir.

4.1 Objectiu de l'anàlisi

L'objectiu d'aquesta implementació és realitzar un estudi i avaluació de les capacitats de reconeixement que tenen els reconeixadors escollits sobre traços de nombres manuscrits a partir del bolígraf digital.

Com es mostrarà a continuació, l'entorn de proves s'ha dut a terme en diferents fases. En un segon punt, s'explicarà la implementació d'una base de dades pròpia i l'ús d'una d'externa.

4.2 Entorn de proves

Es poden apreciar quatre fases en el desenvolupament d'aquest projecte.

4.2.1 Presa de contacte: primeres proves amb el bolígraf digital i l'adquisició de dades

L'objectiu d'aquesta fase inicial és tenir la primera presa de contacte amb el bolígraf digital i l'adquisició de dades d'un formulari mitjançant les llibreries i APIs que disposen Logitech® i Anoto™.

La recollida d'unes primeres dades escrites per diferents usuaris amb el bolígraf digital pot resoldre dos requeriments: l'entrenament necessari pel reconeixedor Heloise Hse i una col·lecció de mostres que seran analitzades amb els diferents reconeixadors, dels que s'obtindran els primers resultats.

Per a dita recollida de mostres, es crea un formulari senzill amb el que es demanarà a una sèrie d'usuaris l'escriptura dels nombres compresos entre el 0 i

el 9 un total de tres cops cadascun.

Aquestes mostres adquirides amb el bolígraf es guardaran en una base de dades per poder treballar amb elles amb flexibilitat ja que, com es veurà més endavant, seran avaluades diferents capacitats dels reconeixadors com els reconeixements dependent i independent d'escriptor.

L'estructura d'aquesta base de dades (anomenem-la "*pròpia*") serà descrita a la secció 4.3, dedicada a tal efecte.

La idea és que aquestes mostres s'utilitzin, en un principi, per ser analitzades seguint la tècnica *one-against-all*, és a dir, es selecciona una mostra com a test i la resta del conjunt serveix per a l'entrenament del reconeixedor Heloise Hse. Tot i que no requereix entrenament, el reconeixedor de Microsoft® també serà avaluat amb les mateixes mostres.

4.2.2 Primera fase: comunicació amb els reconeixadors i primers resultats

En aquesta primera fase es tracta d'obtenir uns primers resultats dels reconeixadors sobre les mostres recollides.

Treballar amb el primer reconeixedor no és problema ja que tant el bolígraf, eines d'AnotoTM i el propi reconeixedor disposen d'una API de desenvolupament Java. En canvi, per poder enviar la informació dels traços, per ser analitzats amb el reconeixedor de Microsoft® i esperar una resposta, es necessita una via de comunicació. Per aquest motiu s'ha fet dues aplicacions diferents, una per recollir i gestionar les dades del bolígraf en Java i una altra, en C#, per executar el reconeixement d'aquestes dades.

Davant l'objectiu de transmetre les dades als reconeixadors i obtenir els re-

sultats, se'ns plantegen diferents punts de decisió:

1. *¿Com passem les dades als reconeixadors?*

Existeix un projecte provinent de la Universitat d'Stanford[21] que fan servir una comunicació entre les aplicacions desenvolupades mitjançant *streaming* de la localització dels arxius XML i la resolució de l'anàlisi per un port.

El reconeixedor de Microsoft® té funcionalitats per treballar amb dades en format XML, per tant es disposaran les dades de cada àrea del formulari amb traços escrits i seran convertides en un XML molt senzill identificant els punts i els traços que formen els nombres.

El programa client, que gestiona la recepció de les dades provinents del bolígraf, executa el servidor amb el reconeixedor de Microsoft® i aquest retorna, per la mateixa via esmentada anteriorment, el resultat obtingut.

En el cas del reconeixedor Heloise Hse, analitzar i obtenir resultats dels traços és més senzill ja que està fet en Java i només s'ha d'importar les classes necessàries i incloure el codi a l'entorn de proves. El seu entrenament, en canvi, sí que necessita ser creat d'alguna forma i carregat pel propi reconeixedor. Així doncs, es crearà un fitxer XML on es guardaran les dades d'entrenament en una carpeta d'on el reconeixedor les llegirà.

2. *¿Com crear els diferents XML necessaris?*

Degut a que tots dos reconeixadors fan servir fitxers XML, per propòsits diferents però amb el mateix contingut, es podria simplificar la creació dels diferents arxius XML. Per això, s'ha implementat un paquet en Java (anomenat *dataRecollection*) que gestiona la recollida d'informació de la base de dades i creació dels fitxers XML.

Aquest, s'explicarà més detingudament a la secció 7.4 dels Annexes, figura 7.2.

3. *¿Quina és la resposta dels reconeixadors?*

Tots dos reconeixadors donen com a mínim una alternativa de reconeixement davant un conjunt de traços, i fins a un màxim de tres alternatives en cas de que la precisió d'una única solució sigui “dubtosa”.

En cas que els reconeixadors retornin més d'una alternativa, aquestes vénen ordenades de major a menor probabilitat de precisió. Esmentar que, malauradament, no es pot obtenir el valor d'aquesta probabilitat per treballar amb ella.

4. *Valoració d'aquests primers resultats*

Els resultats han estat avaluats amb una matriu de confusió dels nombres esperats i obtinguts, així com el percentatge d'encert de cadascun dels reconeixadors.

4.2.3 Segona fase: implementació amb base de dades externa i avaluació Precision-Recall

Ja s'ha produït el primer contacte i una petita avaluació amb els reconeixadors a partir de les mostres de nombres recollides. Però, el fet d'entrenar un reconeixador amb una base de dades pròpia i la dificultat d'obtenir prou mostres, dona sensació de poca independència.

Per aquest motiu es decideix ampliar l'entorn d'avaluació i resoldre certs dubtes, per exemple:

1. *¿Quina base de dades fer servir per entrenar?*

La independència i la heterogeneïtat d'usuaris i mostres és molt important

en els conjunts d'entrenament per als reconeixadors de patrons en general. Existeixen bases de dades externes que garantitzen aquestes característiques.

En concret es farà servir una proporcionada mitjançant un enllaç a la Associació Internacional de Reconeixement de Patrons (International Association for Pattern Recognition - IAPR), Technical Committee 11. S'aprofundirà més sobre ella a la secció que s'hi dedica (4.3).

2. *¿Com s'entrena l'Heloise Hse amb aquesta base de dades externa?*

Per poder entrenar el reconeixedor Heloise Hse amb aquesta base de dades externa s'ha de comprovar que les dades i el format en el que estan emmagatzemades son correctes i concorden amb les dades que el reconeixedor espera rebre.

Les dades es troben en un arxiu de text pla en parells de nombres corresponents a coordenades X i Y, i a l'hora en conjunts de diferents traços i caràcters. Tot separat per etiquetes. Per tant, l'adquisició, tractat i us d'aquestes dades és fàcilment automatitzable amb una computadora.

Tot i que aquests estan creats sobre un espai de 500 per 500 punts i el bolígraf obté una quantitat de punts major o menor segons l'àrea especificada amb les eines d'AnotoTM, no és necessari modificar els seus valor, degut a que el propi reconeixedor té un mètode per estandarditzar els traços abans de fer-los servir per entrenar o be per reconèixer.

L'únic punt a tenir en compte és que el reconeixedor i el bolígraf tenen l'eix Y de l'àrea a l'invers de l'eix de les mostres d'entrenament. Per tant, invertint cadascun dels valors de l'eix Y a les mostres d'entrenament, és suficient i proporciona un entrenament correcte.

Degut a que la base de dades externa no disposa de marques de temps, aquesta característica serà substituïda per zeros a l'entrenament. El re-

coneixement serà igualment correcte ja que els punts dels diferents traços tenen un ordre vectorial, o on-line, al igual que les mostres adquirides amb el bolígraf digital.

3. *Mètode d'avaluació complementari a les matrius de confusió*

En àmbits de reconeixement de patrons és molt recorrent una parella de característiques del reconeixement. La recuperació i la rellevància d'una informació.

La combinació d'aquestes dues característiques d'una informació es mesura amb les probabilitats **Precision** i **Recall** [22].

- **Precision:** mesura la qualitat d'un sistema de recuperació d'informació en termes de la habilitat del sistema per obtenir només elements correctes.
- **Recall:** mesura la efectivitat d'un sistema per obtenir els elements rellevants.

Gràcies a la inclusió d'una base de dades externa com a entrenament del reconeixedor Heloise Hse, podem pensar que aquest serà un bon entrenament i contribuirà a bons resultats. Es diferiran els entrenaments en dos tipus:

- **Independent d'escriptor:** l'Heloise Hse, entrenat únicament amb la base de dades externa, ens assegura una independència d'usuari i obtenció de bons resultats.
- **Dependent d'escriptor:** l'Heloise Hse, entrenat amb la base de dades externa, propícia l'obtenció de bons resultats. A més, el fet d'afegir a l'entrenament traços dels mateixos escriptors, proporciona un reconeixement personalitzat de nombres.

També dir que l'avaluació d'aquests resultats millora per a futures preses de decisió amb la incorporació de la parella Precision-Recall. Tant el reconeixedor

de Microsoft® com les variants del Heloise Hse seran avaluats amb les matrius de confusió i una taula Precision-Recall.

El flux d'execució i dades en aquesta segona fase es pot veure a la secció 7.4.2 figura 7.3.

4.2.4 Tercera fase: combinació de reconeixadors

A la fase anterior s'ha desenvolupat una avaluació *precision* i *recall* dels diferents reconeixadors amb les diferents configuracions, tant dependent com independent d'escriptor. Aquests valors s'han inclòs en taules (una per cada reconeixedor) que acumulen les capacitats d'aquests en funció dels nombres analitzats i el nombre d'alternatives extretes. Les taules (7.1, 7.2, 7.3) es poden veure al capítol “Annexes”.

Ja que les alternatives ofertes pels dos reconeixadors, donat un conjunt de traços, estan en ordre de major a menor probabilitat segons cada reconeixedor, el nou reconeixedor mixte escollirà les probabilitats *precision* i *recall* pertanyents a cada nombre corresponent a cada posició dins les taules Precision-Recall.

La multiplicació de les probabilitats *precision* i *recall* donen com a resultat una altre probabilitat (que s'anomenarà d'**encert** [$Prob.Encert = Prob.Precision \cdot Prob.Recall$]). Aquesta probabilitat resulta la compensació entre les capacitats de recuperació i precisió dels reconeixadors. Es dona així un alt valor probabilístic d'encert a les alternatives dels reconeixadors amb característiques més compensades i altes, i per tant millors.

D'aquesta manera, s'aconsegueix una llista de probabilitats d'encert conjunta amb valors únics i ordenats de major a menor. Dels quals, per defecte, s'escull el resultat que té una probabilitat d'encert major.

4.3 Bases de dades

Per la realització d'aquest projecte s'ha disposat de dos bases de dades. Una **pròpia**, que serveix per analitzar i un entrenament dependent d'escriptor; i una **externa**, que ens garanteix un bon entrenament.

4.3.1 Base de dades pròpia

A cadascun d'un total de 20 usuaris se'ls ha demanat escriure la seqüència de nombres del 0 al 9 amb tres iteracions de cadascun d'ells. En total 600 mostres.

El motiu és poder recollir suficients dades per crear una petita primera base de dades per l'entrenament, i alhora, tenir prou mostres per poder analitzar cadascuna d'elles de forma dependent i independent d'escriptor. Es pot veure l'estructura utilitzada a la figura 4.1 i s'especifiquen els atributs al capítol “Annexes”, secció 7.3.

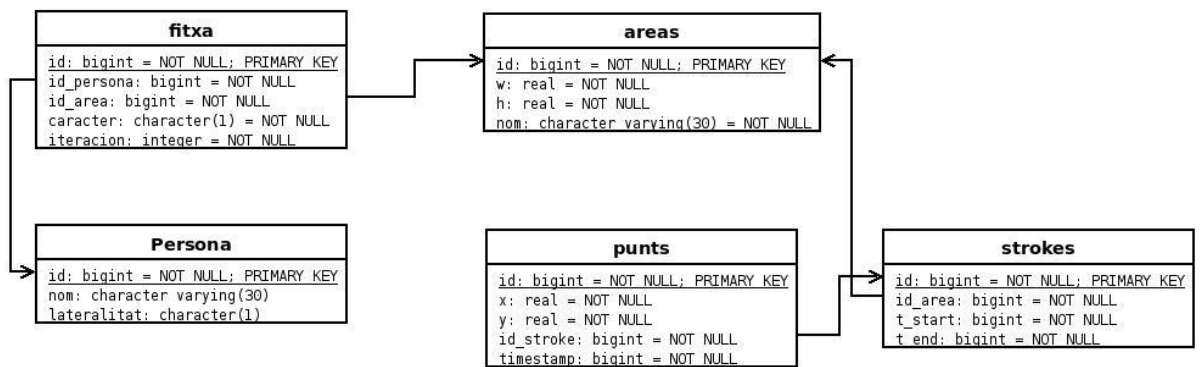


Figura 4.1: Mapa relacional de la base de dades que serveix per emmagatzemar les dades extretes del bolígraf digital amb els diferents traços dels usuaris que han participat.

La base de dades creada ha tingut una finalitat molt concreta: emmagatzemar

les dades que s'han recollit amb el bolígraf digital per poder treballar amb elles i analitzar els diferents reconeixadors.

4.3.2 Base de dades externa

La necessitat d'un bon entrenament del reconeixador Holise Hse és la clau per a que funcioni de forma òptima. És necessari una base de dades creada per la funcionalitat concreta de reconèixer on-line els nombres de qualsevol tipus d'escriptor. Això implica un gran nombre de mostres, heterogeneïtat i "control de qualitat" sobre els diferents escriptors que han participat en el recull dels models.

Unipen Foundation Inc. [23] és una entitat que s'ha encarregat de recollir mostres i caràcters de diferents tipus per entrenar i provar els diferents reconeixadors que es desenvolupen, o programari que necessiti bases de dades de caràcters on-line. S'han encarregat també de crear un sistema per guardar aquestes dades d'una forma simple, portable i que no faci servir gaires recursos del sistema. Es tracta del format estandar UNIPEN.

Aquest sistema, mitjançant etiquetes, especifica línia a línia les diferents característiques que té la base de dades. Aquestes etiquetes es diferencien per començar amb un punt i el nom de l'etiqueta amb majúscules seguit dels diferents paràmetres. Cada traç de les mostres s'especifica per una parella .PEN_DOWN - .PEN_UP i les coordenades X Y on s'ha escrit cada punt del traç. També trobem altres etiquetes que especifiquen el caràcter al que pertanyen els traços, nombre de traços per caràcter, o quins son els caràcters que s'inclouen al fitxer de la base de dades que s'està llegint.

Les bases de dades d'aquesta fundació son molt complertes. Tenen a disposició des de nombres, lletres en minúscula o majúscula i, fins i tot, conjunts de paraules

i textos. Tots manuscrits on-line. i son a l'abast de tothom en un CD que la fundació ven al preu de \$110.

Per a aquest projecte s'ha estimat més utilitzar una base de dades que, tot i no ser la esmentada degut al preu, utilitza la mateixa estructura estàndard. També està recomanada per la Associació Internacional per al Reconeixement de Patrons (International Association for Pattern Recognition - IAPR[24]), Technical Committee 11. Més concretament, ha estat proporcionada per Ethem Alpaydin i Fevzi. Alimoglu del *Department of Computer Engineering, Bogazici University*. Aquesta és d'us lliure i gratuïta. Tot i que només recull les dades sobre nombres manuscrits en format on-line, és just el que es necessita.

L'arxiu utilitzat per entrenar el reconeixedor Heloise Hse disposa d'un total de 7494 nombres escrits per 30 escriptors i aproximadament 250 nombres cadascun de forma aleatòria.

Capítol 5

Avaluació dels sistemes de reconeixement

Per a fer l'avaluació del sistema de recull i anàlisis de dades obtingudes amb el bolígraf digital, els resultats analitzats seran els provinents dels respectius reconeixadors. Al llarg d'anar concretant les fases de desenvolupament, els resultats analitzats passaran a tenir un caire més enfocat a la dependència i independència d'escriptor dels reconeixadors per separat i de la combinació dels mateixos.

5.1 Tècniques d'anàlisi de resultats

El concepte a analitzar, al llarg del projecte, ha passat de ser cadascun dels dos reconeixadors escollits per separat, a analitzar la forma de complementar i combinar aquests per tal d'aconseguir els millors resultats en solucions dependents i independents de l'escriptor.

Les tècniques utilitzades per realitzar aquest anàlisi han estat les matrius de confusió i les probabilitats Precision-Recall.

- **Matrius de confusió:**[25]

Es tracta d'una matriu on cada valor correspon al número de cops que el reconeixedor a identificat un patró dins les diferents possibilitats.

La posició d'aquest valor dins de la matriu ve determinat per la fila, segons el valor corresponent al resultat obtingut pel reconeixedor; i per columna, corresponent a la posició del valor del patró que es tracta realment.

Una matriu de confusió amb un 100% d'encert, és a dir que totes les mostres han estat reconegudes correctament, es visualitzaria com una matriu diagonal.

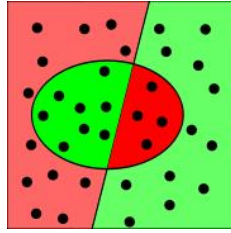
- **Precision-Recall:**[22, 26, 27]

El binomi *Precision-Recall* son unes probabilitats molt utilitzades en reconeixement de patrons. La seva finalitat és la mesura de la qualitat d'encert d'un sistema en termes de la habilitat per només incloure elements rellevants en els resultats (*Precision*) i la efectivitat del sistema en obtenir només els elements rellevants (*Recall*).

Ve definit per les fórmules de la figura 5.1, on *rel* és el nombre d'elements rellevants dins un conjunt de mostres i *ret* el nombre d'elements obtinguts dins el mateix conjunt.

Table : Retrieval Matrix

	Relevant	Non-Relevant	TOTAL
Retrieved	$ ret \cap rel $ ■	$ ret \cap \overline{rel} $ ■	$ ret $
Not Retrieved	$ \overline{ret} \cap rel $ ■	$ \overline{ret} \cap \overline{rel} $ ■	$ \overline{ret} $
TOTAL	$ rel $	$ \overline{rel} $	$ tot $



$$P = \frac{|ret \cap rel|}{|ret|}$$

$$R = \frac{|ret \cap rel|}{|rel|}$$

Figura 5.1: A les imatges es defineixen les probabilitats *Precision*(P) i *Recall*(R) al avaluar uns resultats en termes de reconeixement de patrons. Com es pot veure, donat un conjunt de mostres, aquestes es poden dividir entre rellevants i no rellevant, i entre obtingudes o no obtingudes.

5.2 Avaluació dels resultats

En aquesta secció s'analitzaran els resultats obtinguts al llarg de l'avaluació dels reconeixadors. Aquest anàlisi es dividirà en tres apartats corresponents a cada fase de la implementació.

5.2.1 Reconeixedors de Microsoft® i Heloise Hse amb base de dades pròpia

A la primera fase s'han analitzat les mostres adquirides amb el bolígraf digital amb cadascun dels reconeixadors. En aquest cas, l'Heloise Hse ha estat entrenat amb la mateixa base de dades pròpia d'on provenen les mostres a ser analitzades. S'ha separat de l'entrenament, però, la mostra a analitzar per l'entorn dependent

d'escriptor i totes les mostres de l'usuari per l'entorn independent d'escriptor.

A la taula 5.1, es mostren els valors que representen els encerts obtinguts respecte la quantitat de nombres recuperats dins el subgrup de nombres entre les alternatives (Recuperats), i respecte el total de les 600 mostres analitzades (Total).

# alternatives	Encerts/Respecte	MSreco	HHreco - D.E.	HHreco - I.E.
1 alt.	Recuperats	418/424(98.58%)	378/548(68.98%)	316/536(58.95%)
1 alt.	Total	418/600(69.67%)	378/600(63.00%)	316/600(52.66%)
2 alt.	Recuperats	118/123(95.93%)	23/31(74.19%)	20/44(45.45%)
2 alt.	Total	118/600(19.66%)	23/600(3.83%)	20/600(3.33%)
3 alt.	Recuperats	50/53(94.34%)	14/20(70.00%)	17/20(85.00%)
3 alt.	Total	50/600(8.33%)	14/600(2.33%)	17/600(2.83%)

Taula 5.1: *Comparativa entre els resultats dels reconeixadors en la primera fase d'avaluació.*

De tots els resultats obtinguts, els més importants son els pertanyents als valors encertats com a primera alternativa. Aquests demostren la capacitat, no només d'encert, sinó també d'obtenció dels reconeixadors.

Es poden destacar els alts valors del reconeixedor proporcionat per Microsoft® respecte les alternatives recuperades (entre **99%** i **94%** aproximadament).

Tot i així, no és suficient respecte les 600 mostres analitzades (**69.67%** obtinguts i encertats pel reconeixedor respecte el total de mostres). Des d'aquest punt de vista, els resultats s'assemblen més als obtinguts amb les opcions dependent (**63.00%**) i independent (**52.66%**) d'escriptor configurades al reconeixedor Heloise Hse.

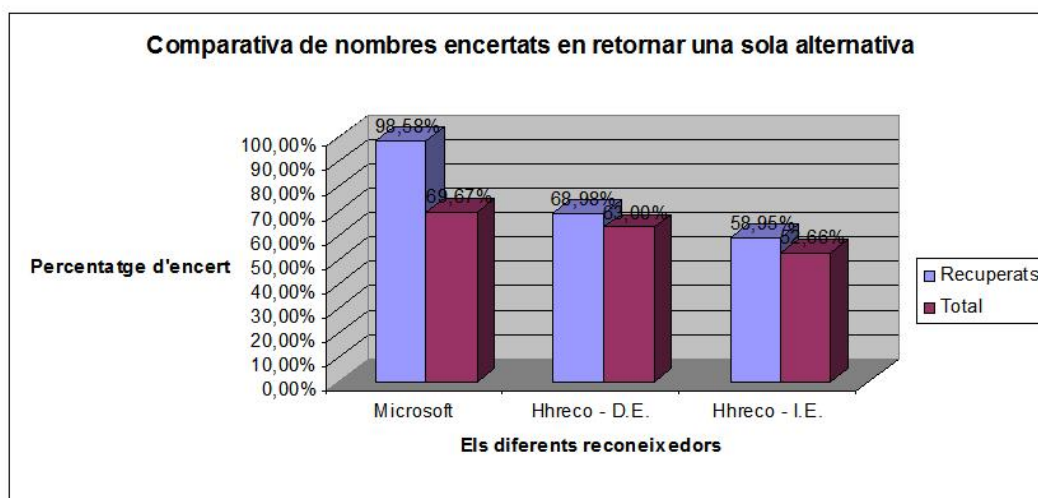


Figura 5.2: *Comparativa de nombres encertats en retornar una sola alternativa respecte els recuperats per cada reconeixedor i respecte el total de mostres analitzades.*

5.2.2 Reconeixedor de Microsoft® i Heloise amb base de dades externa

Després d'obtenir els primers resultats amb els reconeixedor implementats a la primera fase, és natural pensar que el reconeixedor de Microsoft® no és gaire millorable (per si sol) pel fet que no té cap opció d'entrenament específic. En canvi, l'Heloise Hse ha estat entrenat amb les mateixes mostres amb que s'ha fet l'avaluació. En total 600 mostres. Per això, en aquesta fase l'entrenament es proporcionarà amb la base de dades externa.

Els resultats (taules 5.2 i 5.3, i figures 7.5 - 7.7) mostren un lleuger increment de la capacitat d'encert del reconeixedor Heloise Hse amb una base de dades externa, tant per la configuració dependent com per la independent d'escriptor. Així, els nous resultats obtinguts, s'aproximen als resultats del reconeixedor de Microsoft® respecte el total de mostres analitzades.

# alternatives	Encerts/Respecte	HHreco - D.E. - B.D. Pròpia	HHreco - D.E. - B.D. Externa
1 alt.	Recuperats	378/548(68.98%)	419/568(73.77%)
1 alt.	Total	378/600(63.00%)	419/600(69.83%)
2 alt.	Recuperats	23/31(74.19%)	11/24(45.83%)
2 alt.	Total	23/600(3.83%)	11/600(1.83%)
3 alt.	Recuperats	14/20(70.00%)	6/8(75.00%)
3 alt.	Total	14/600(2.33%)	6/600(1.00%)

Taula 5.2: *Comparativa entre els resultats de la primera fase i els obtinguts amb la millora en el Heloise Hse dependent d'escriptor.*

# alternatives	Encerts/Respecte	HHreco - I.E. - B.D. Pròpia	HHreco - I.E. - B.D. Externa
1 alt.	Recuperats	316/536(58.95%)	368/543(67.77%)
1 alt.	Total	316/600(52.66%)	368/600(61.33%)
2 alt.	Recuperats	20/44(45.45%)	24/45(53.33%)
2 alt.	Total	20/600(3.33%)	24/600(4.00%)
3 alt.	Recuperats	17/20(85.00%)	7/10(70.00%)
3 alt.	Total	17/600(2.83%)	7/600(1.17%)

Taula 5.3: *Comparativa entre els resultats de la primera fase i els obtinguts amb la millora en el Heloise Hse independent d'escriptor.*

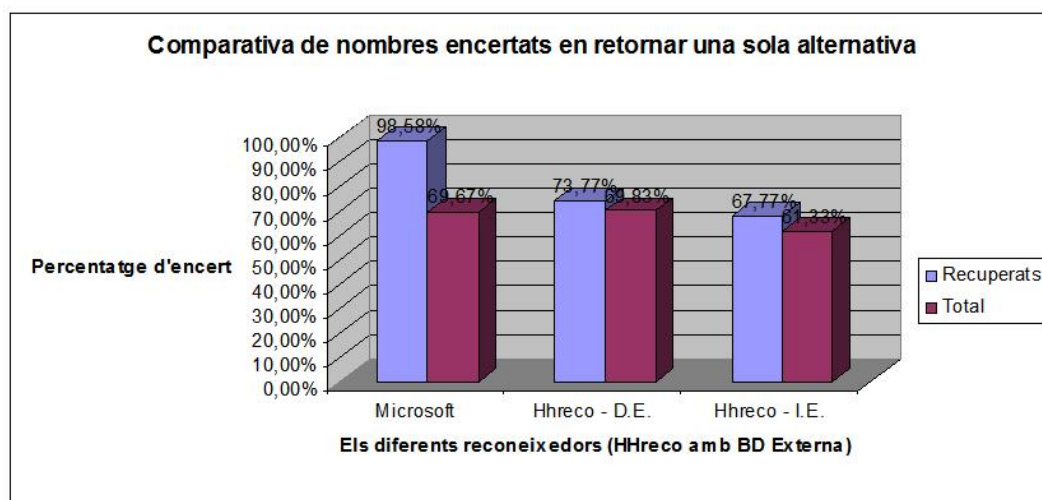


Figura 5.3: *Comparativa de nombres encertats en retornar una sola alternativa respecte els recuperats per cada reconeixedor i respecte el total de mostres analitzades.*

En tot cas, els dos reconeixadors deixen sense especificar aproximadament al 30% o 40% del total de les mostres analitzades (repartides en els subgrups de 2 o

3 alternatives possibles). Aquests percentatges no es consideren bons, ja que la finalitat és reconèixer el màxim nombre de mostres analitzades de forma segura.

Considerant que, després del nou entrenament del reconeixedor Heloise Hse, tots dos reconeixadors estan ben entrenats, es decideix fer un anàlisi dels mateixos amb un canvi: en comptes de separar els resultats dels reconeixements en diferents grups d'alternatives, es considerarà bona la primera alternativa de cada grup. Es a dir, el resultat que cada reconeixedor determinaria com a millor d'entre les diferents alternatives.

Reconeixedors	Microsoft	Heloise Hse(I.E.)	Heloise Hse(D.E.)
Percentatge d'encert	568/600(94.67%)	385/600(64.17%)	428/600(71.33%)

Taula 5.4: *Percentatges d'encert dels diferents reconeixadors prenent com a bona la primera de les diferents alternatives que retornen.*

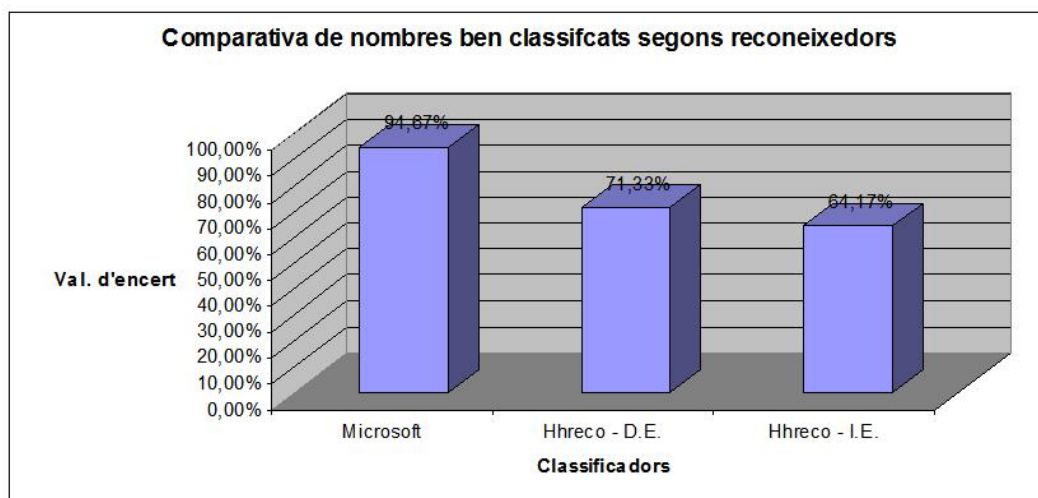


Figura 5.4: *En aquesta imatge es mostra com el reconeixedor de Microsoft® dona millors resultats si es dona com a encertada la primera alternativa que aquest considera més probable d'encert.*

Com es pot veure a la taula i figura 5.4, el reconeixedor Heloise Hse no té la mateixa eficiència reconeixent nombres com el de Microsoft® a l'hora de determinar una única solució. Però es continuarà al llarg de l'estudi amb la configuració

dels reconeixadors per extreure d'una a tres alternatives, ja que, en aquest cas, sí sembla haver una equitat entre els reconeixements i possibilitats d'aplicar alguna millora que augmenti els bons resultats obtinguts fins aquest moment.

5.2.3 Combinació dels reconeixadors

En aquesta última fase es veurà que, amb la combinació dels diferents reconeixadors, es pot augmentar al percentatge d'encert. La solució està en fer-ho juntament amb les taules Precision-Recall creades a la segona fase. Taules (7.1, 7.2 i 7.3).

Per posar un exemple pràctic, si volem reconèixer el número 4, passarem els traços corresponents a aquest número als reconeixadors i aquests retornen unes possibles alternatives de reconeixement: MS (4, 5) i HH (2, 4, 6). En total tindríem 5 alternatives per relacionar amb el número que volem identificar (4:PR=97.78% , 5:PR=37.35% , 2:PR=39.37% , 4:PR=29.63% , 6:PR=33.33%).

Si es sumen les probabilitats de les diferents alternatives (en aquest cas las del nombre 4) i es divideixen per la suma de totes elles, s'obtindrà una probabilitat d'encert del 53.66% de que el nombre 4 és l'esperat i totes les altres opcions seran desestimades. Veure més concretament la taula 5.5 a partir de les taules Precision-Recall 7.1 i 7.2, i la figura 7.4 al capítol “Annexes”.

Nombre	Probabilitat P*R	Probabilitat P*R'	Probabilitat d'encert(%)
(MS 1) “4”	97.78%	127.41	53.66%
(MS 2) “5”	37.35%	37.35	15.73%
(HH 1) “2”	39.37%	39.37	16.58%
(HH 2) “4”	29.63%	–	–
(HH 3) “6”	33.33%	33.33	14.04%

Taula 5.5: *Taula que mostra un exemple del funcionament del reconeixedor mixte avaluant els traços d'un nombre 4.*

Procedint amb aquest mètode s'obtindrà un reconeixement de nombres que

destacarà lo bo i dolent de cada reconeixedor i discretitzarà els resultats per obtenir l'encert més probable. Com a resultat, la taula 5.6.

Reconeixedors	Mixte - I.E.	Mixte - D.E.
Percentatge d'encert	570/600(95.00%)	573/600(95.50%)

Taula 5.6: *Percentatges d'encert amb el reconeixedor mixte.*

Es poden veure a les figures 5.5 i 5.6 els resultats comparatius entre els reconeixedor (configuracions dependent i independent d'escriptor respectivament) on s'avalua la capacitat d'encert en anàlisis d'única alternativa emesa.

En la configuració dependent d'escriptor, observar els resultats obtinguts del reconeixedor mixte (600 mostres analitzades, 573 mostres correctament identificades: **95.50% d'efectivitat**). (Matriu de Confusió de la figura 7.9). Tot i que el reconeixedor de Microsoft® no es considera dependent d'escriptor, s'afegeix per a simple comparació d'efectivitat.

D'altra banda, en la configuració independent d'escriptor, els resultats també son destacables (600 mostres analitzades, 570 mostres correctament identificades: **95.00% d'efectivitat**). (Matriu de Confusió de la figura 7.8).

Finalment, igual que a la fase anterior, s'han comparat els resultats obtinguts només avaluant els resultats en cas que el reconeixedor hagués d'escollir una solució d'entre les seves alternatives. Veure figura 5.7.

Com es pot veure, ara sí s'obté un guany en el reconeixement dels nombres. Això es degut a la implicació de les capacitats de reconeixement de cada reconeixedor. S'han tingut en compte les febleses i bondats de cadascun a l'hora de reconèixer els patrons i s'ha dissolt el nivell de "dubte" davant les diferents alternatives emeses pels mateixos reconeixedor.

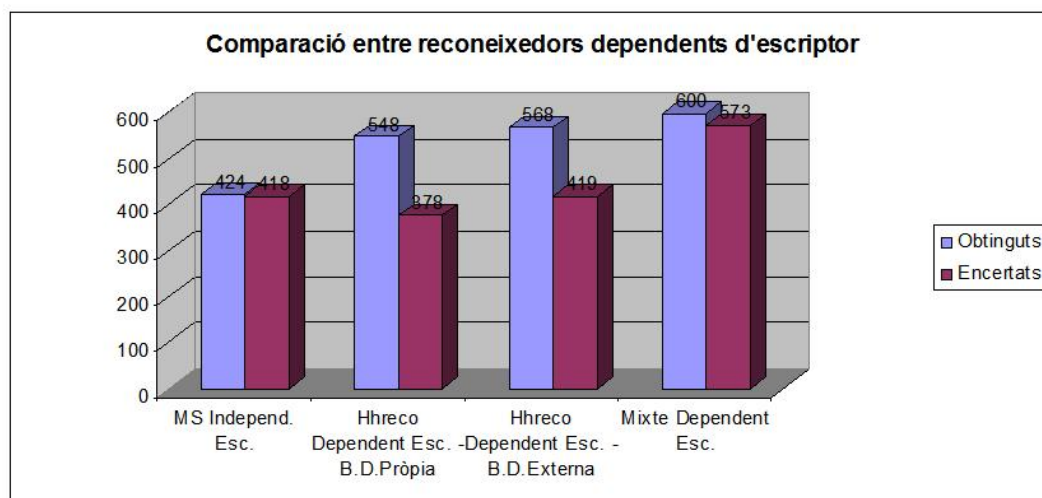


Figura 5.5: *Resultats obtinguts en anàlisi d'única alternativa emesa pels diferents reconeixadors i configuració **dependents** d'escriptor.*

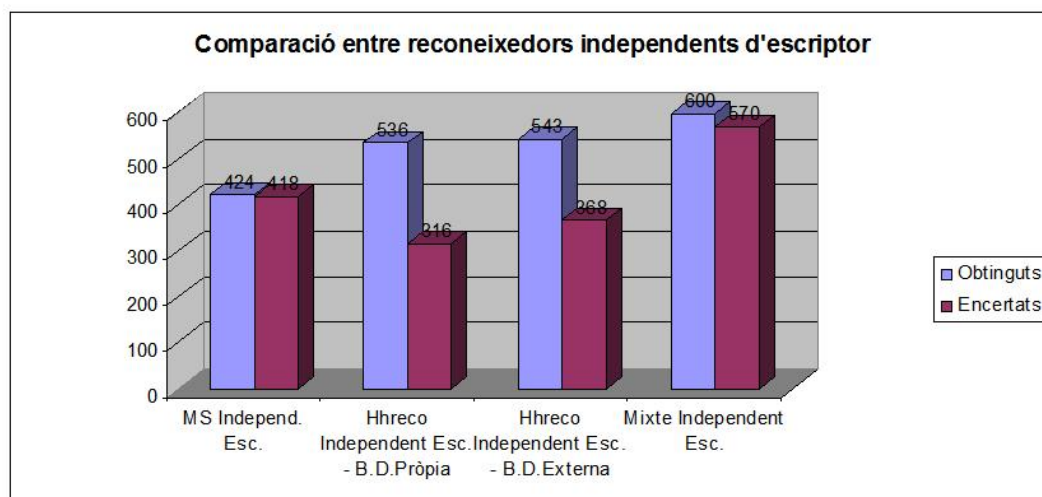


Figura 5.6: *Resultats obtinguts en anàlisi d'única alternativa emesa pels diferents reconeixadors i configuració **independents** d'escriptor*

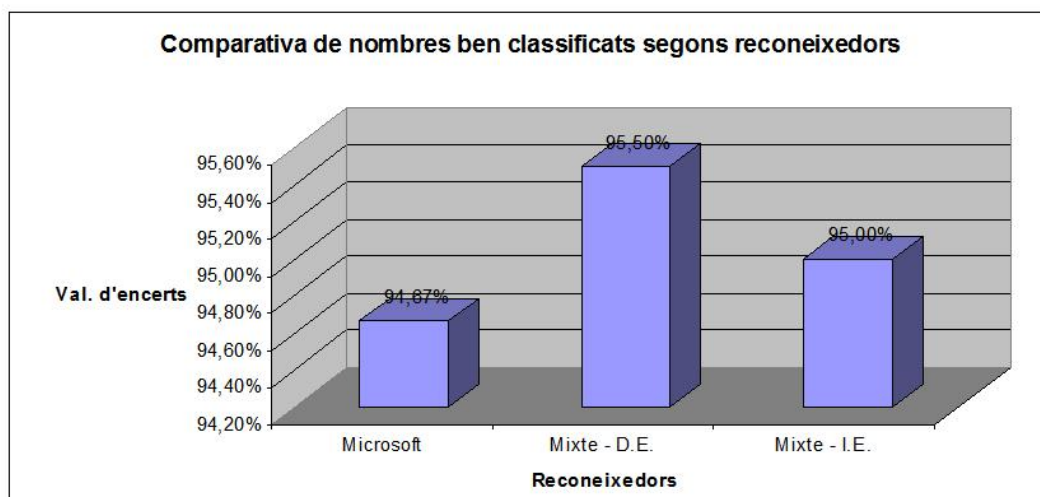


Figura 5.7: Resultats comparant el reconeixedor de Microsoft® i la solució mixte donant com a solució bona la més probable segons cadascun.

Capítol 6

Conclusions

Finalment, aquest capítol està destinat a explicar les conclusions obtingudes de les proves realitzades. Es comentarà l'assoliment dels objectius i les possibles línies que es podrien determinar en un futur.

6.1 Objectius assolits

Recordar que l'objectiu d'aquest projecte és l'anàlisi de reconeixadors (i cercar una possibilitat de millora) de nombres manuscrits amb un bolígraf digital. Els resultats seran determinants a l'hora d'incloure la capacitat d'un reconeixedor en l'administració de formularis digitalitzats.

Després d'analitzar els reconeixadors per separat, amb diferents combinacions i, finalment, una combinació d'ambdós per tal d'extreure el millor de cadascun; es considera que els resultats obtinguts son prou bons per fer possible l'aplicació de reconeixement a la solució final.

La decisió de quin serà el reconeixedor escollit, o quina configuració s'utilitzarà, es deixa en mans de l'empresa usufructuària. Això és degut a que, tant amb el reconeixedor de Microsoft® com amb la solució mixte de tots dos reconeixadors (tant dependent com independent d'escriptor), els resultats han estat molt similars i prou bons per fer servir qualsevol de les opcions.

Segons les avaluacions realitzades, els reconeixadors mixtes tenen uns millors resultats simbòlics (**rec. mixte dependent d'escriptor 95,50% d'encert, rec. mixte independent d'escriptor 95,00% d'encert**) que el reconeixedor de Microsoft® (**94,67% d'encert**).

Degut a la doble infraestructura de reconeixement que comporta avaluar uns traços mitjançant ambdós reconeixadors, i per tant un petit descens del rendiment del sistema, la proposta d'aplicació és implantar només el sistema de reconeixement de Microsoft® per al reconeixement de nombres.

En el cas que ambdós reconeixadors fossin necessaris dins del mateix sistema, per a diferents tasques (reconeixement de textos o nombres més reconeixement de formes o figures), sí que seria adequat combinar les respostes dels reconeixadors

i aprofitar les seves característiques amb les taules Precision-Recall. Tècnica que demostra un augment de prestacions destacant les virtut i minimitzant els defectes de cada reconeixedor.

6.2 Línies futures

L'art del reconeixement de documents no es queda només en l'àmbit numèric. Amb les possibilitats que donen el bolígraf digital de Logitech® i els formularis i eines d'Anoto™, el mon de la indústria pot veure's molt beneficiat.

Una línia futura evident és no quedar-se només amb el reconeixement de nombres, sinó ampliar al conjunt de caràcters existents i, mitjançant segmentació o reconeixement de paraules i expressions, reconèixer textos sencers com és el cas del projecte STATE de la Universitat Jaume I de Castelló[28, 29].

Una altra línia de reconeixement és el de símbols o figures. És interessant al mon industrial dissenyar un sistema que, al dibuixar esbossos a mà, aquest reconegués i informatitzés amb figures estandarditzades i mides a escala un resultat final. Aquesta via podria ser viable amb el reconeixedor, analitzat en aquest document, Heloise Hse.

Per tal d'estandarditzar el transport de les dades i incorporar la capacitat de desenvolupar solucions amb tinta digital sobre internet, un punt d'estudi interessant és la utilització i benefici que pot aportar l'estandar InkML [30], un format XML estandar de dades creat expressament per solucions d'escriptura amb traços digitals.

Capítol 7

Annexes

7.1 Requeriments del sistema

Tant AnotoTM com Logitech[®] només proporcionen controladors i entorn de desenvolupament per entorns Microsoft[®] Windows. En aquest punt ens ajustarem a les versions del sistema operatiu i requeriments que s'aconsellen a les especificacions.

Els requeriments del sistema son:

- PC compatible amb IBM[®].
- Windows[®] 2000, Windows[®] XP Home i Professional.
- Pentium[®] III a 500 MHz (es recomana a 700 MHz o superior).
- 256 MB de RAM (es recomana 512 MB).
- Port USB disponible.
- Resolució de pantalla mínima de 800 x 600 i 256 colors.

- Microsoft® .NET Framework 1.1 (inclòs al CD adjunt) o posterior.
- Microsoft® Internet Explorer 5.01 (inclòs al CD adjunt) o posterior.
- Aproximadament 500 MB d'espai lliure al disc.

7.2 Flux d'execució i manipulació amb solucions AnotoTM

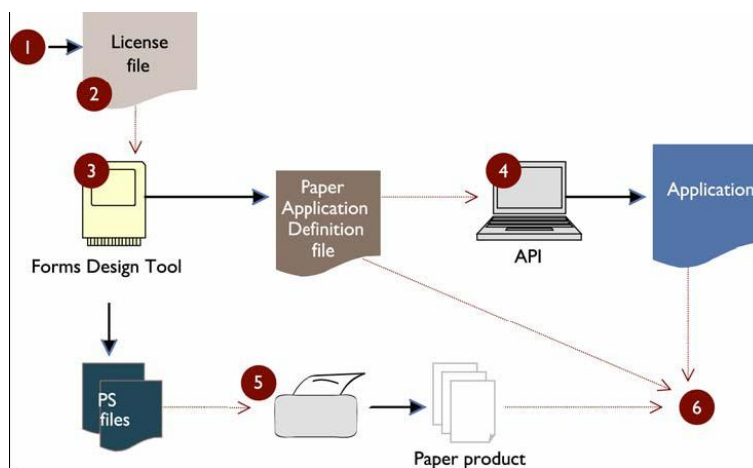


Figura 7.1: Es mostra el flux d'execució i manipulació dels components de la solució d'AnotoTM

1. Adquisició de llicència.
2. Aplicació de la llicència i especificació de les àrees d'usuari.
3. Creació del formulari i l'arxiu *.PAD.
4. Programació de la manipulació de les dades adquirides amb el bolígraf.
5. Impressió del formulari creat per a ser omplert per un usuari.

6. Un cop omplert els formularis, tractar la informació recollida amb el bolígraf (i emmagatzemada en un fitxer *.PGC) segons les especificacions programades al punt 4.

7.3 Base de dades pròpia

Especificació dels atributs de les taules:

- **fitxa:**

Taula creada com a punt d'unió entre els caràcters recollits, les àrees d'escriptures a les que pertanyen i l'usuari que els ha escrit.

- **id: Clau primària.** Identificador únic dins la taula. Autoincremental.
- **id_persona:** *Clau referencial.* Identificador de l'usuari.
- **id_area:** *Clau referencial.* Identificador de l'àrea on s'ha escrit.
- **caracter:** Valor amb certesa que té l'àrea escrita. Nombres del 0 al 9.
- **iteracion:** El nombre de la iteració al qual pertany aquell àrea. (Primera, segona o tercera iteració)

- **persona:**

Taula on s'emmagatzemen els usuaris que participen en l'anàlisi.

- **id: Clau primària.** Identificador únic dins la taula. Autoincremental.
- **nom:** Nom de l'usuari.
- **lateralitat:** Indicació si l'usuari és esquerrà o dretà. Atribut no utilitzat. Pensat per anàlisis posteriors en línies futures.

- **area:**

Taula on s'emmagatzemen les dades de l'àrea on s'ha escrit els traços al paper intel·ligent.

- **id: Clau primària.** Identificador únic dins la taula. Autoincremental.
- **w:** Amplada de l'àrea.
- **h:** Alçada de l'àrea.
- **nom:** Nom atribuït a l'àrea per diferenciar-la de les altres.

- **strokes:**

Taula amb el recull dels traços fets per cada caràcter escrit pels usuaris. Cada àrea pot tenir més d'un traç o *stroke*.

- **id: Clau primària.** Identificador únic dins la taula. Autoincremental.
- **id_area:** *Clau referencial.* Identificador de l'àrea on es localitza aquest traç.
- **t_start:** Marca de temps inicial de tota la seqüència de punts que formen el traç.
- **t_end:** Marca de temps final de tota la seqüència de punts que formen el traç.

- **punts:**

Taula que recull tots els punts amb coordenades X, Y i marques de temps que formen cadascun dels traços.

- **id: Clau primària.** Identificador únic dins la taula. Autoincremental.
- **x:** Coordenada X de la localització del punt.
- **y:** Coordenada Y de la localització del punt.

- **id_stroke**: *Clau referencial*. Identificador del traç al qual pertany aquest punt.
- **timestamp**: Marca de temps en l'instant que s'ha recollit aquest punt.

7.4 Disseny de les diferents fases

7.4.1 Disseny a la primera fase

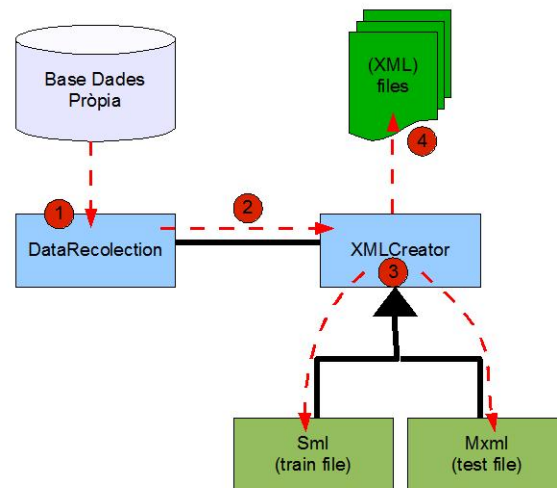


Figura 7.2: Aquest esquema mostra com es creen els diferents fitxers XML que es necessitaran tant per entrenar com testejar els reconeixadors.

1. Elecció de les dades amb les que es vol crear el fitxer XML.
2. Informació al XMLCreator de quines son les dades escollides i el tipus de fitxer que es vol crear.
3. XMLCreator executa els mètodes de les classes Sml i Mxml pre crear l'estructura segons convingui (entrenament pel reconeixedor Heloise Hse o test pel reconeixedor de Microsoft®.)

4. Creació dels fitxers amb les dades pertinents i emmagatzemament on s'especifiqui per cada tipus de fitxer.

7.4.2 Disseny a la segona fase

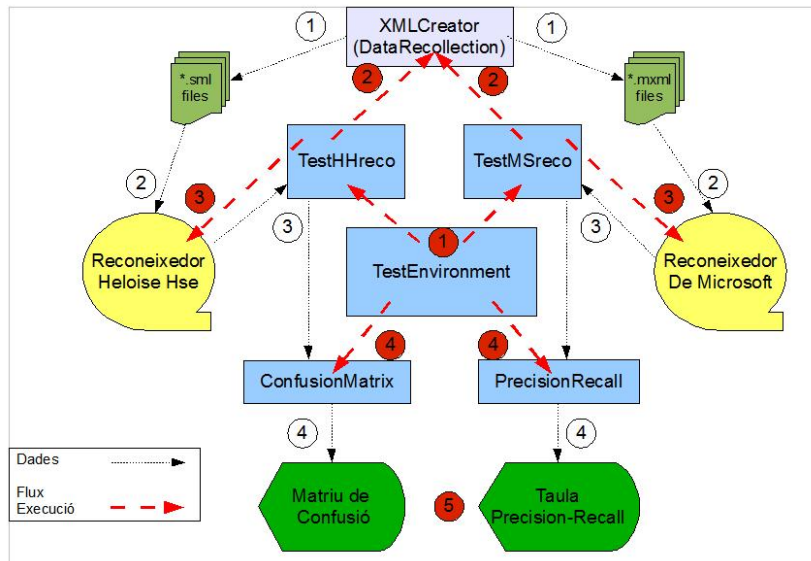


Figura 7.3: Es mostren tant el flux d'execució (nombres i fletxes en vermell) com el camí de les dades (nombres i fletxes en blanc i negre) dins l'entorn de test ideat per l'avaluació dels reconeixadors.

1. A la classe principal (TestEnvironment) es troba la funció que, successivament, recollirà cadascun dels nombres a analitzar i li indicarà a les classes TestHHreco i TestMSreco.
2. Aquestes cridaran al DataRecollection. Aquest, està encarregat de proporcionar els arxius XML de les dades d'entrenament i test per a cada reconeixedor (segons les necessitats de cadascun) [dades (1)]. Els disposarà on s'especifiqui per a que posteriorment siguin recollits i tractats pels reconeixadors segons convingui (entrenament o test) [dades (2)].

3. Els reconeixadors, instanciats també per les classes TestHHreco i TestMSreco, retornen els valors de les alternatives extretes després de l'anàlisi de cada nombre.
4. Aquestes alternatives seran rebudes per la classe principal (TestEnvironment) i preparades per a ser avaluades mitjançant una matriu de confusió i una taula de probabilitats Precision-Recall [dades (3)].
5. Les classes ConfusionMatrix i PrecisionRecall extreuran els resultats de l'avaluació per pantalla [dades (4)].

7.4.3 Disseny a la tercera fase

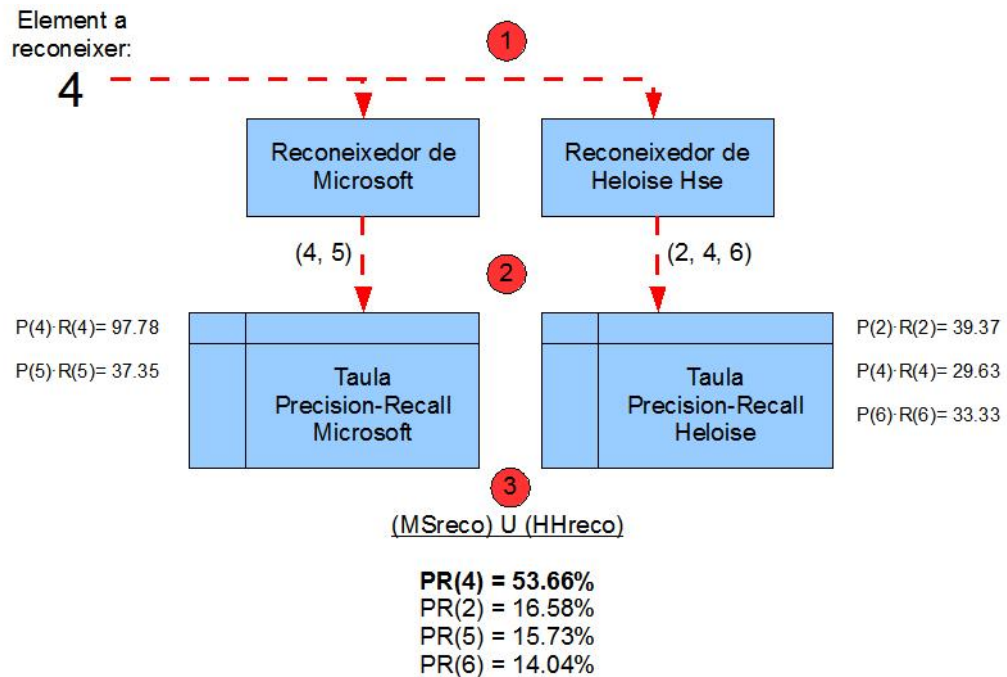


Figura 7.4: En aquest esquema es pot veure els diferents passos que impliquen la millora i barreja dels resultats dels reconeixadors.

1. Avaluació dels traços per part dels reconeixadors.

2. Obtenció de les diferents alternatives proposades, per cada reconeixedor, donat una conjunt de traços. Consulta dels valors Precision-Recall, associats a cada alternativa, dins les taules Precision-Recall de cada reconeixedor.
3. Suma dels valors Precision-Recall de les alternatives coincidents i obtenció del percentatge que representen respecte el total.

7.5 Taules Precision-Recall

A continuació es mostren les taules *Precision-Recall* utilitzades com a característiques dels reconeixadors en la implementació del reconeixedor mixte.

Símbol:	"0"	"1"	"2"	"3"	"4"	"5"	"6"	"7"	"8"	"9"
Recall - 1 alt.(%)	55/55(1.00)	48/49(0.98)	46/46(1.00)	39/39(1.00)	44/45(0.98)	32/35(0.91)	55/55(1.00)	40/40(1.00)	31/32(0.97)	28/28(1.00)
Precision - 1 alt.(%)	55/57(0.96)	48/51(0.94)	46/46(1.00)	39/39(1.00)	44/44(1.00)	32/32(1.00)	55/55(1.00)	40/41(0.96)	31/31(1.00)	28/28(1.00)
Recall - 2 alt.(%)	3/3(1.00)	9/9(1.00)	12/12(1.00)	17/17(1.00)	4/7(0.57)	20/21(0.95)	3/3(1.00)	14/14(1.00)	17/18(0.94)	19/19(1.00)
Precision - 2 alt.(%)	3/15(0.20)	9/24(0.36)	12/20(0.6)	17/32(0.53)	4/26(0.15)	20/51(0.39)	3/8(0.38)	14/27(0.52)	17/20(0.85)	19/23(0.83)
Recall - 3 alt.(%)	1/2(0.50)	2/2(1.00)	2/2(1.00)	4/4(1.00)	8/8(1.00)	3/4(0.75)	2/2(1.00)	6/6(1.00)	9/10(0.90)	13/13(1.00)
Precision - 3 alt.(%)	1/10(0.10)	2/19(0.11)	2/17(0.12)	4/15(0.27)	8/21(0.38)	3/16(0.19)	2/10(0.2)	6/10(0.60)	9/19(0.47)	13/22(0.59)

Taula 7.1: *Taula Precision-Recall associada al reconeixedor de Microsoft® en subconjunts de 1, 2 o 3 alternatives obtingudes. Marcats en negreta els valors de l'exemple.*

Símbol:	"0"	"1"	"2"	"3"	"4"	"5"	"6"	"7"	"8"	"9"
Recall - 1 alt.(%)	58/60(0.97)	34/56(0.61)	27/48(0.56)	40/54(0.74)	35/53(0.66)	31/54(0.57)	33/57(0.58)	41/57(0.72)	43/57(0.75)	26/47(0.55)
Precision - 1 alt.(%)	58/58(1.00)	34/54(0.63)	27/38(0.71)	40/57(0.70)	35/51(0.69)	31/53(0.58)	33/42(0.79)	41/62(0.66)	43/57(0.75)	26/71(0.37)
Recall - 2 alt.(%)	0/0(0.00)	2/4(0.50)	3/8(0.38)	0/4(0.00)	4/6(0.67)	5/6(0.83)	2/2(1.00)	0/3(0.00)	3/3(1.00)	5/9(0.56)
Precision - 2 alt.(%)	0/0(0.00)	2/9(0.22)	3/11(0.27)	0/10(0.00)	4/9(0.44)	5/11(0.45)	2/12(0.17)	0/6(0.00)	3/10(0.30)	5/12(0.42)
Recall - 3 alt.(%)	0/0(0.00)	0/0(0.00)	3/4(0.75)	1/2(0.50)	1/1(1.00)	0/0(0.00)	1/1(1.00)	0/0(0.00)	0/0(0.00)	1/2(0.50)
Precision - 3 alt.(%)	0/0(0.00)	0/2(0.00)	3/5(0.60)	1/2(0.50)	1/5(0.20)	0/7(0.00)	1/3(0.33)	0/1(0.00)	0/0(0.00)	1/5(0.20)

Taula 7.2: *Taula Precision-Recall associada al reconeixedor de Heloise Hse (Independent d'escriptor + B.D. Externa) en subconjunts de 1, 2 o 3 alternatives obtingudes. Marcats en negreta els valors de l'exemple.*

Símbol:	"0"	"1"	"2"	"3"	"4"	"5"	"6"	"7"	"8"	"9"
Recall - 1 alt.(%)	58/60(0.97)	39/54(0.72)	35/55(0.64)	44/57(0.77)	44/59(0.75)	39/57(0.68)	38/56(0.68)	42/58(0.72)	46/56(0.82)	34/56(0.61)
Precision - 1 alt.(%)	58/58(1.00)	39/52(0.75)	35/44(0.80)	44/55(0.80)	44/65(0.68)	39/57(0.68)	38/51(0.75)	42/58(0.72)	46/62(0.74)	34/66(0.52)
Recall - 2 alt.(%)	0/0(0.00)	0/4(0.00)	2/4(0.50)	1/2(0.50)	0/0(0.00)	2/3(0.67)	1/3(0.33)	1/2(0.50)	2/3(0.67)	2/3(0.67)
Precision - 2 alt.(%)	0/1(0.00)	0/5(0.00)	2/5(0.40)	1/4(0.25)	0/3(0.00)	2/7(0.29)	1/4(0.25)	1/4(0.25)	2/3(0.67)	2/12(0.17)
Recall - 3 alt.(%)	0/0(0.00)	0/2(0.00)	1/1(1.00)	1/1(1.00)	1/1(1.00)	0/0(0.00)	1/1(1.00)	0/0(0.00)	1/1(1.00)	1/1(1.00)
Precision - 3 alt.(%)	0/0(0.00)	0/1(0.00)	1/3(0.33)	1/3(0.33)	1/3(0.33)	0/3(0.00)	1/3(0.33)	0/2(0.00)	1/2(0.50)	1/4(0.25)

Taula 7.3: *Taula Precision-Recall associada al reconeixedor de Heloise Hse (Dependent d'escriptor + B.D. Externa) en subconjunts de 1, 2 o 3 alternatives obtingudes.*

7.6 Matrius de Confusió dels resultats

Tot seguit es mostren les matrius de confusió principals dels diferents reconeixadors, així com les corresponents als resultats obtinguts amb el reconeixedor mixte.

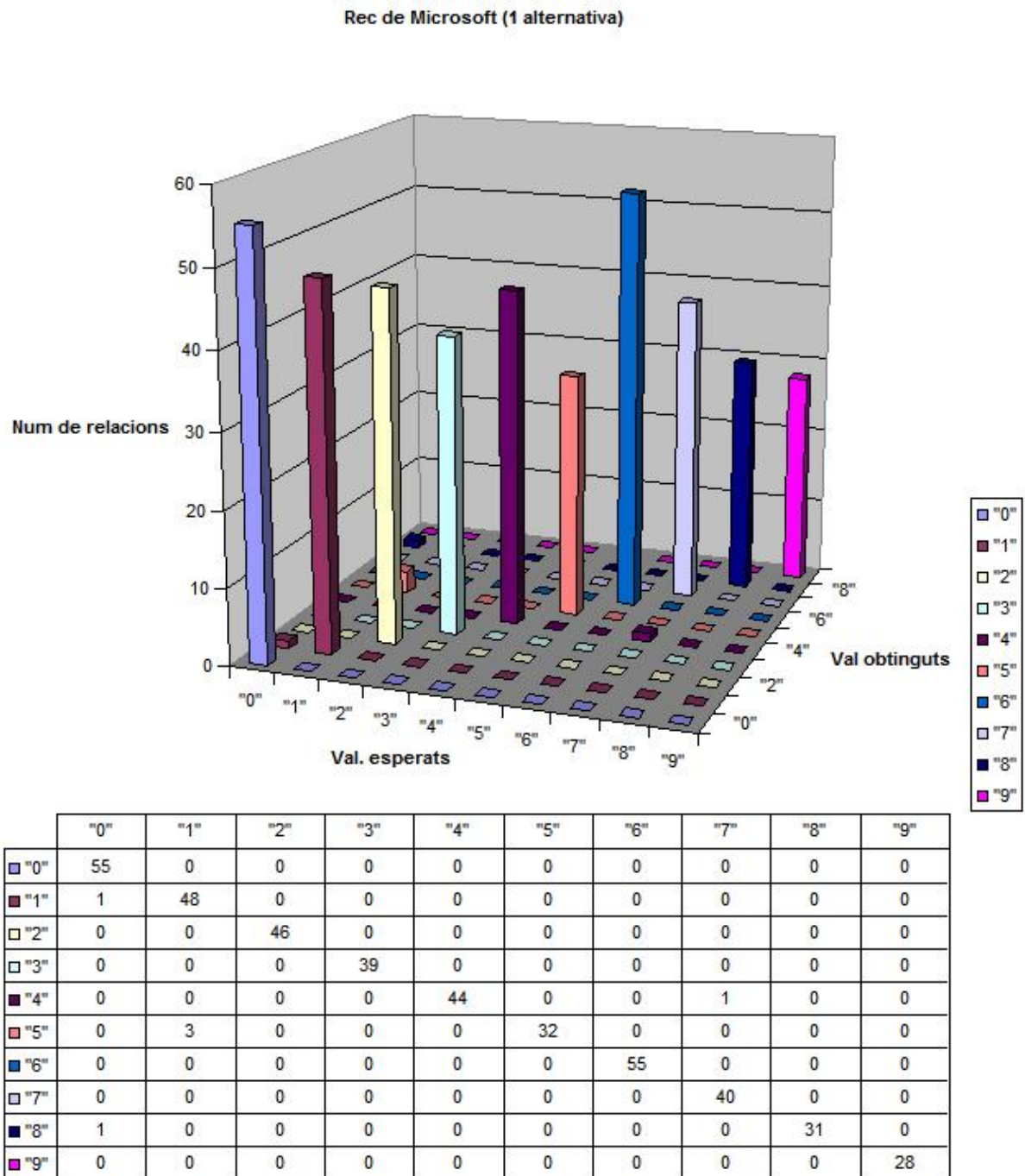


Figura 7.5: *Matriu de confusió dels resultats obtinguts com a única alternativa amb el reconeixedor de Microsoft®.*

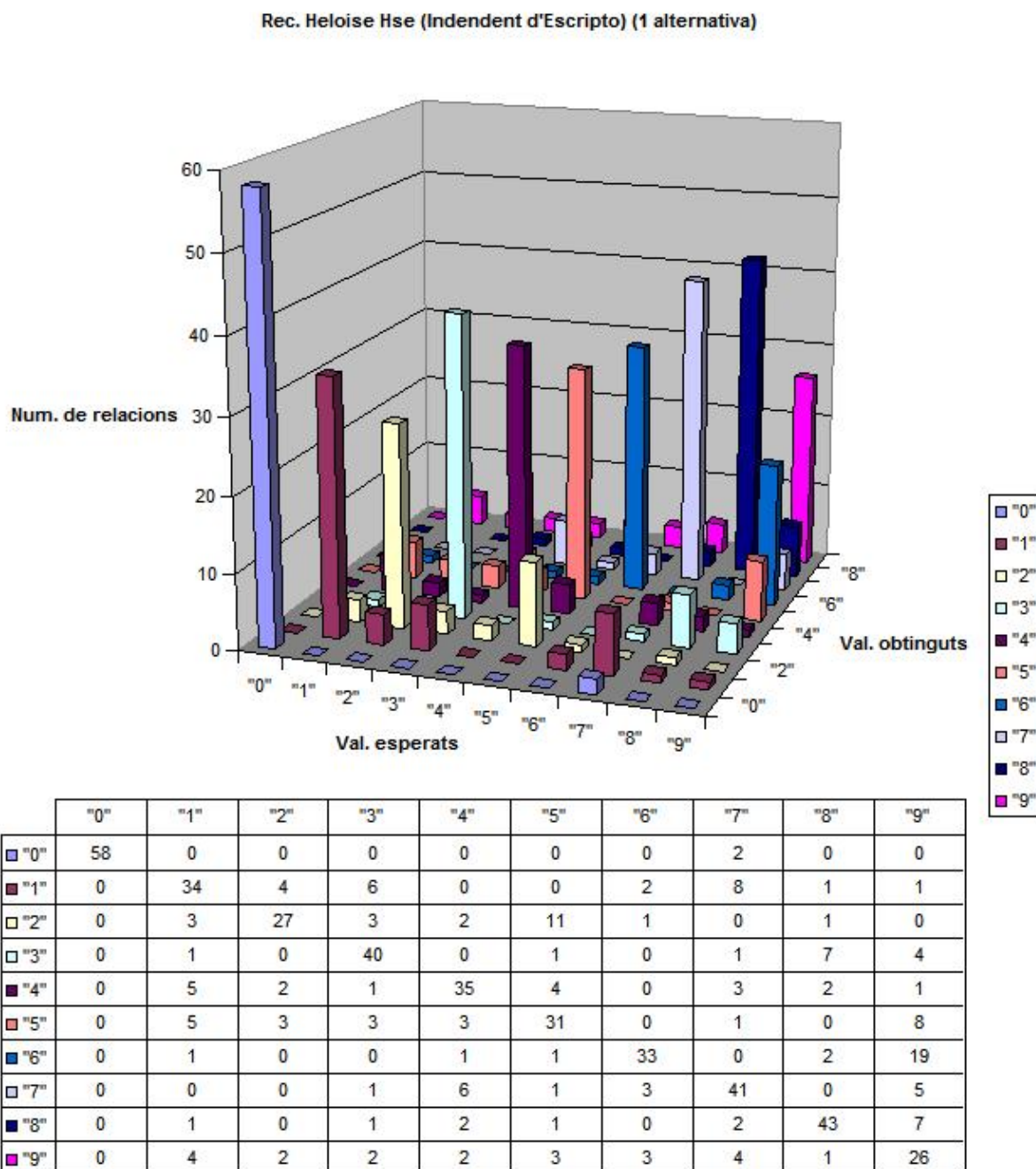


Figura 7.6: *Matriu de confusió dels resultats obtinguts com a única alternativa amb el reconeixedor Heloise (Independent d'escriptor - B.D Externa).*

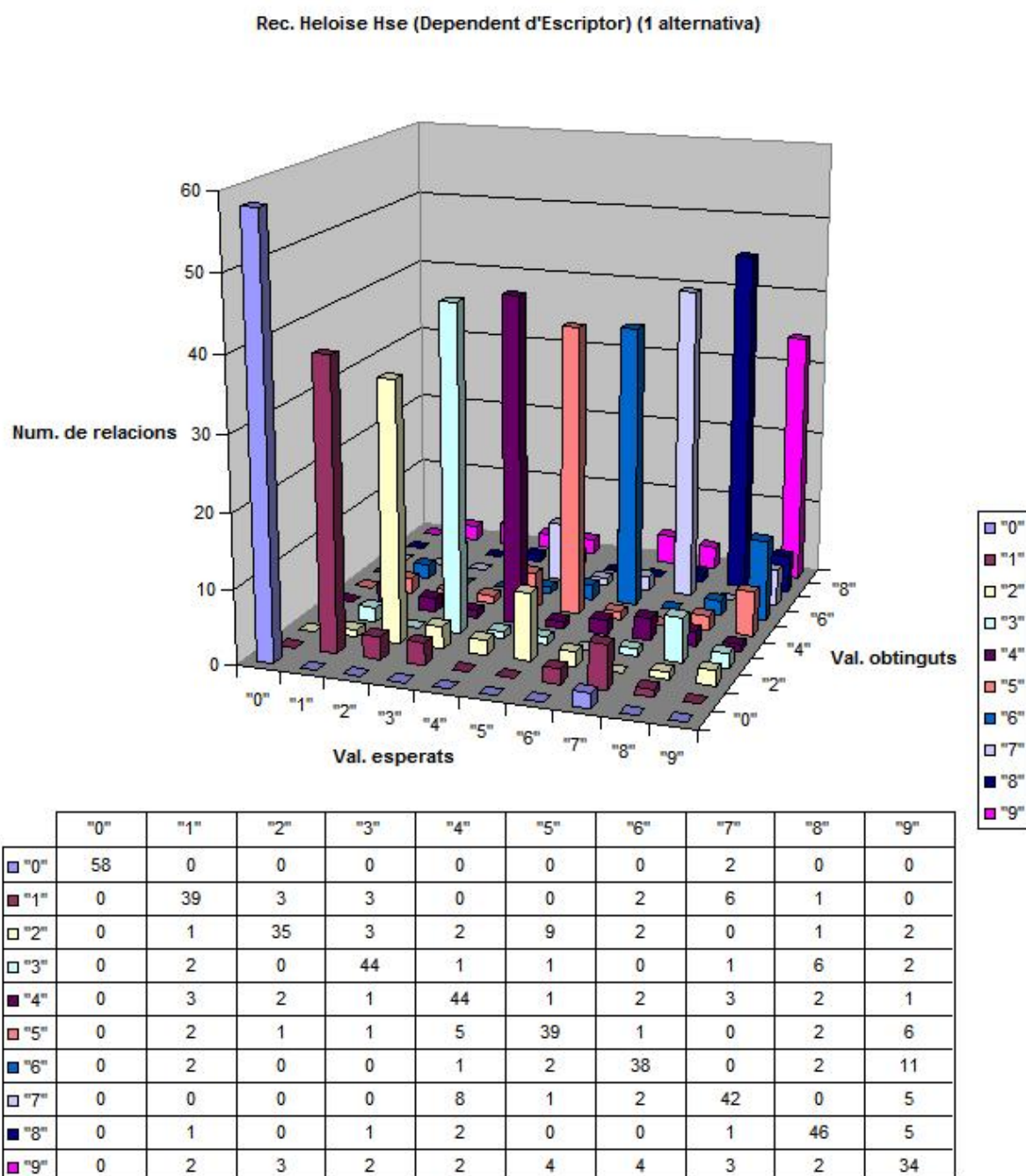


Figura 7.7: *Matriu de confusió dels resultats obtinguts com a única alternativa amb el reconeixedor Heloise (Dependent d'escriptor - B.D Externa).*

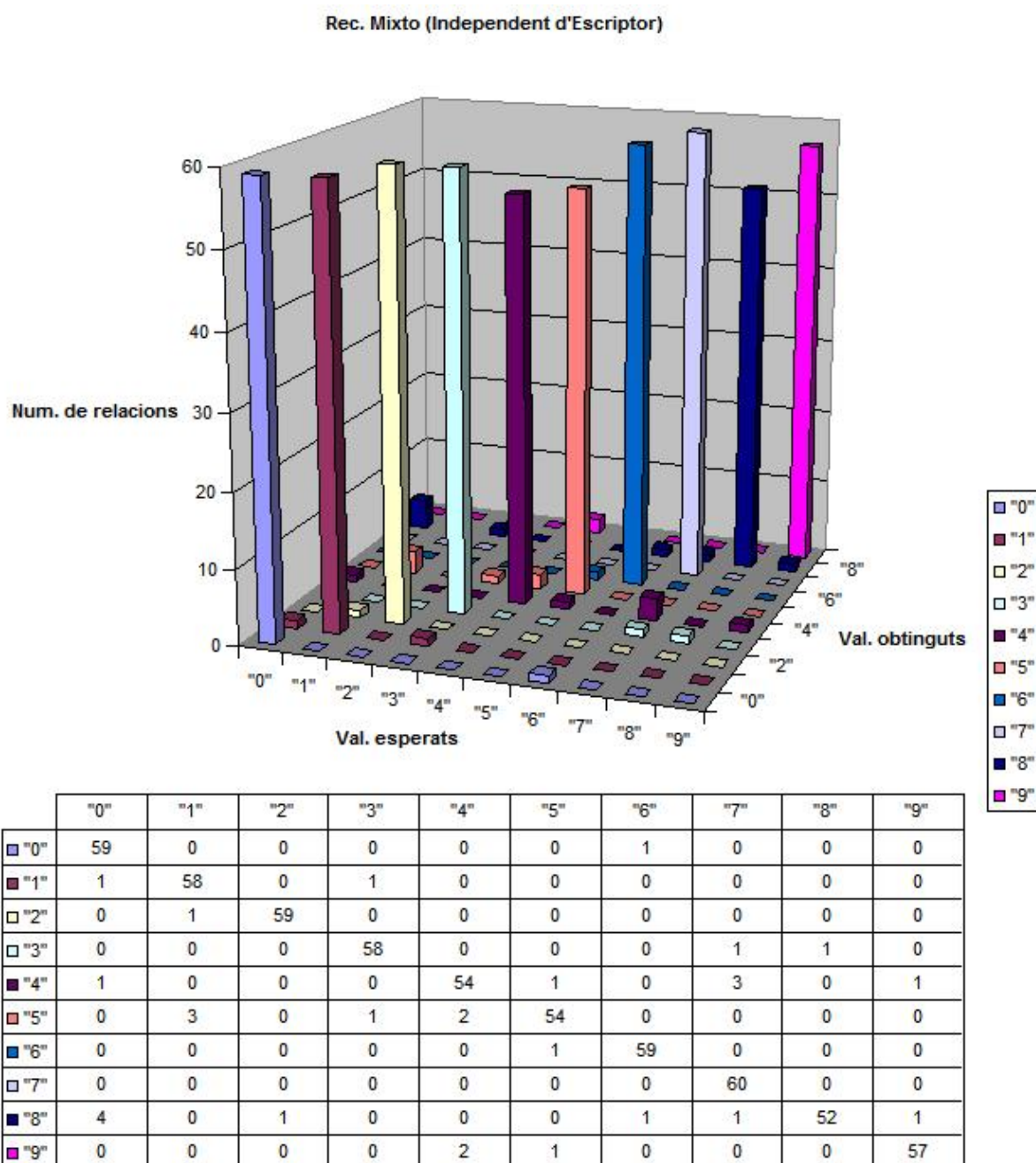


Figura 7.8: Matriu de confusió dels resultats obtinguts com a única alternativa amb el reconeixedor desenvolupat mixte (Independent d'escriptor).

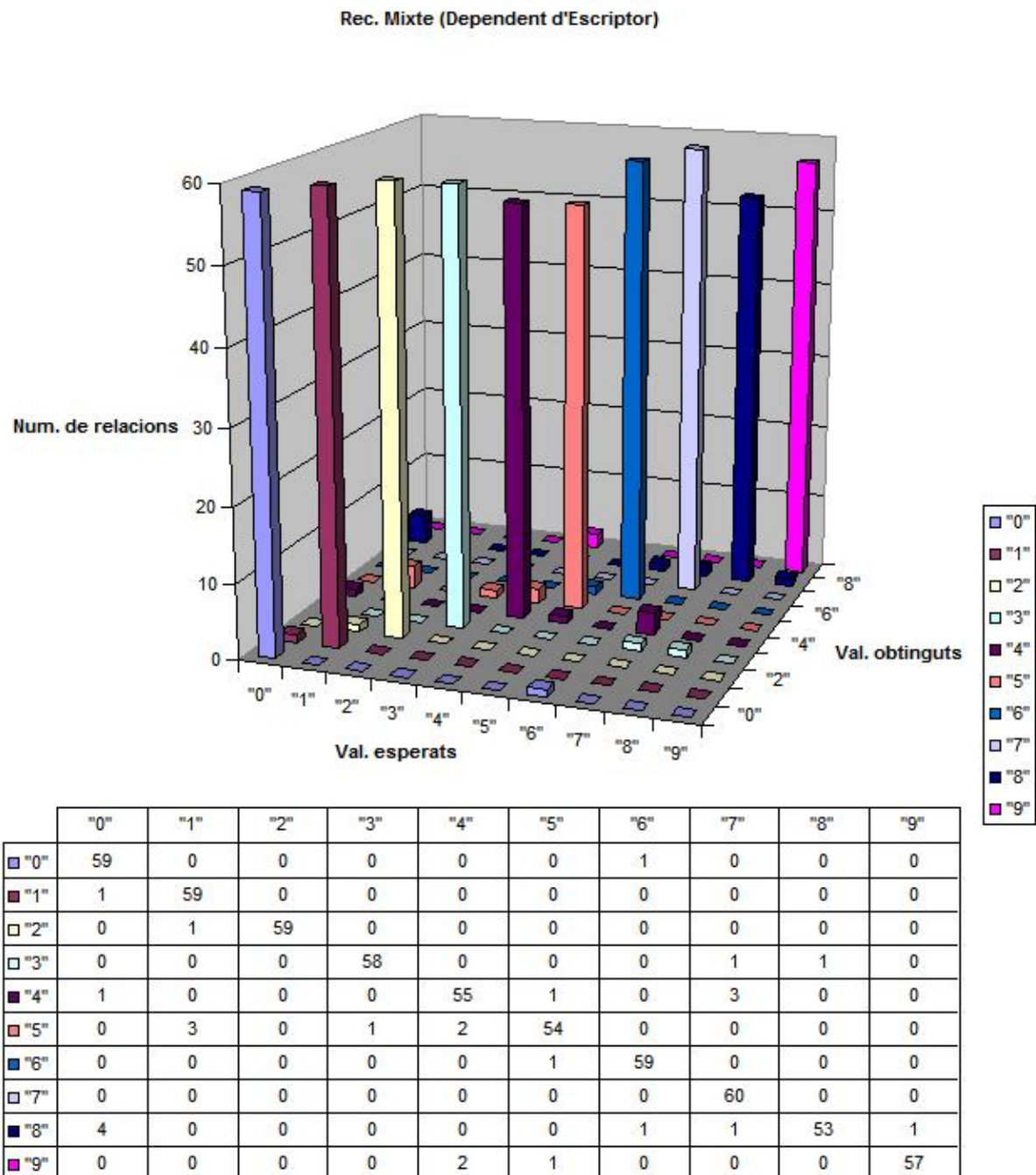


Figura 7.9: *Matriu de confusió dels resultats obtinguts com a única alternativa amb el reconeixedor desenvolupat mixte (Dependent d'escriptor).*

Bibliografia

- [1] Anonymous, “Optical character recognition” -
http://en.wikipedia.org/wiki/Optical_character_recognition
- [2] Scantron®, “¿Qué es OMR, ICR y OCR?” -
<http://espanol.scantron.com/sistemas/comparacion.htm>
- [3] Anonymous, “Handwriting Recognition” -
http://en.wikipedia.org/wiki/Handwriting_recognition
- [4] Stefan Helkvist, “On-line character recognition on small hand-held terminals using elastic matching” - Stockholm 1999
www.hellkvist.org/software/xmerlin/
- [5] Logitech®, “IO™2 Digital Pen” -
http://www.logitech.com/index.cfm/mice_pointers/digital_pen/devices/408&cl=ES,ES
- [6] Anoto®, *<http://www.anoto.com>*
- [7] Microsoft®, “Microsoft Windows XP Tablet PC Edition Software Development Kit 1.7” -
<http://www.microsoft.com/downloads/details.aspx?FamilyID=b46d4b83-a821-40bc-aa85-c9ee3d6e9699&displayLang=en>
- [8] Microsoft®, “Microsoft Windows XP Tablet PC Edition 2005 Recognizer Pack” -

<http://www.microsoft.com/downloads/details.aspx?FamilyID=080184dd-5e92-4464-b907-10762e9f918b&DisplayLang=en>

- [9] Heloise Hse, “HHreco Software Download Page” - (Electrical Engineering and Computer Sciences) University of California at Berkeley -
<http://embedded.eecs.berkeley.edu/research/hhreco/>
- [10] Michael Levin, “CellWriter: Grid-Entry Handwriting Recognition” - (Undergraduate Research Opportunities Program) University of Minnesota
- [11] “ tesseract-ocr: An OCR Engine that was developed at HP Labs between 1985 and 1995... and now at Google.” - (Undergraduate Research Opportunities Program) University of Minnesota -
<http://code.google.com/p/tesseract-ocr/>
- [12] “Wikipedia: Optical Character Recognition” -
http://en.wikipedia.org/wiki/Optical_character_recognition/
- [13] “Wikipedia: Hidden Markov Model” -
http://en.wikipedia.org/wiki/Hidden_Markov_model
- [14] “Wikipedia: K-Nearest Neighbor Algorithm” -
http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm/
- [15] “Wikipedia: Neural Network” -
http://en.wikipedia.org/wiki/Neural_network
- [16] “Wikipedia: Support Vector Machine” -
http://en.wikipedia.org/wiki/Support_vector_machine
- [17] Francesco Camastra, “A SVM-based cursive character recognizer” -
(Department of Applied Science) University of Parthenope.
- [18] Heloise Hse and A. Richard Newton, “Sketched Symbol Recognition using Zernike Moments” -

(Department of Electrical Engineering and Computer Sciences) University of California at Berkeley.

- [19] “Wikipedia: Minimum Distance” -
http://en.wikipedia.org/wiki/Minimum_distance
- [20] “Development guide for services enabled by anoto functionality” -
(Components d’AnotoTM - Inclòs al paquet SDK d’AnotoTM)
- [21] “Paper Applications Toolkit” -
(Stanford HCI Group) Universitat d’Stanford. -
<http://hci.stanford.edu/research/paper/>
- [22] Marçal Rusiñol i Josep Lladós, “Precision, Recall and other performance evaluation measures.” -
(Centre de Visió per Computador - Pattern Recognition and Document Analysis Group) Universitat Autònoma de Barcelona.
- [23] “Unipen Project - Data and Benchmark for Handwriting Recognition” -
<http://www.unipen.org/>
<http://unipen.nici.ru.nl/>
- [24] “IAPR - TC-11: Reading Systems” -
<http://www.iapr-tc11.org/>
- [25] “Wikipedia: Confusion matrix” -
http://en.wikipedia.org/wiki/Confusion_matrix
- [26] “Wikipedia: Precision and Recall” -
http://en.wikipedia.org/wiki/Precision_and_recall
- [27] “Wikipedia: Information Retrieval” -
http://en.wikipedia.org/wiki/Information_retrieval
- [28] “State System - Transcription Tool Speeds Recovery Of Ancient Documents”
-

http://www.scientificblogging.com/news_articles/state_system_transcription_tool_speeds_recovery_ancient_documents

- [29] “Investigadores de la UJI diseñan un sistema de transcripción que permite una recuperación más rápida de documentos antiguos y manuscritos” -

http://www.universia.es/html_estatico/portada/actualidad/noticia_actualidad/param/noticia/baaajb.html

- [30] W3C (World Wide Web Consortium), “Ink Markup Language (InkML)” -

<http://www.w3.org/2002/mmi/ink>

Firmat: Iván Navarro Parreño
Bellaterra, Setembre de 2009

Resum

Aquest projecte consisteix en la realització d'un anàlisi de diferents reconeixadors de caràcters manuscrits, concretament de nombres, per a una possible implantació en la digitalització de formularis en la indústria. Al llarg del document s'estudien dos reconeixadors diferents, concretament l'incorporat al paquet "Tablet PC and Recognition Pack" de Microsoft® i el Heloise Hse, proporcionat per la Universitat de Berkeley a Califòrnia.

Resumen

Este proyecto consiste en la realización de un análisis de diferentes reconocedores de caracteres manuscritos, concretamente de números, para una posible implantación en la digitalización de formularios en la industria. A lo largo del documento se estudian dos reconocedores diferentes, concretamente el incorporado en el paquete "Tablet PC and Recognition Pack" de Microsoft® y el Heloise Hse, proporcionado por la Universidad de Berkeley, California.

Abstract

This project involves an analysis of different manuscript character recognizers, particularly of numbers, for possible implementation in form digitalization of industry. Throughout the paper explores two different recognizers, specifically the incorporated into the "Tablet PC and Recognition Pack" package of Microsoft® and Heloise Hse, provided by the University of California at Berkeley.